RoboFaceMaker: Designing Understandable Facial Expressions to Communicate Robot Intentions

Jan Leusmann [1], Thomas Weber [1], Maximilian Kriebel [1], Chao Wang [2], Sven Mayer [3,4]

***Abstract*— Robots in social and collaborative settings must clearly communicate their internal states and intentions to foster trust and coordination with humans. Facial expressions are a powerful and intuitive channel for such communication. However, creating dynamic, expressive robot faces remains challenging due to a lack of accessible design tools. We present the RoboFaceMaker, a web-based, keyframe animation system that enables researchers and practitioners to design, refine, and deploy facial animations. We validated our approach through three studies: a usability study (N=16), an expert design study (N=8), and a validation study (N=12). Our results show that our RoboFaceMaker enables users to design understandable animations for robot faces. We found that non-anthropomorphic expressions were better understood than those that mimicked human expressions. We contribute an open-source animation creation and playback tool, and insights into how facial expressions can be used to communicate intentions.**

I. Introduction

Robots are becoming part of everyday life in workplaces [1], healthcare [2], and domestic settings [3]. As they transition from functional machines to social partners, they must communicate internal states and intentions in ways humans can quickly and reliably understand. Breakdowns in this communication reduce trust, hinder coordination, and compromise safety. A challenge in HRI is to design expressive behaviors that respect all the tradeoffs [4]. Facial expressions are a powerful channel, intuitively signaling emotions and intentions [5], [6]. Yet, authoring such expressions remains complex and time-consuming, limiting their reliability as a channel for communicating intention.

Prior research shows that non-verbal communication, like facial expressions, gaze, and gestures, shapes how people perceive and collaborate with robots [7], [8], [9]. Expressive faces increase perceptions of competence, trustworthiness, and engagement [10], [11]. Rendered faces are prevalent because they are flexible and inexpensive, e.g., Promo Bot, Navel Robotics. However, existing tools are often limited to static layouts or fixed, pre-defined animations [12], [13]. Other systems rely on synchronizing animations with speech or emotion recognition, or by using large language models (LLMs) for face generation [14], [15]. Thus, researchers lack accessible tools for iteratively designing and deploying facial expressions, limiting support for interactions using facial animations.

[1]LMU Munich, Munich, Germany {jan.leusmann, thomas.weber}@ifi.lmu.de

[2]Honda Research Institute EU, Offenbach am Main, Germany chao.wang@honda-ri.de

[3]TU Dortmund University, Dortmund, Germany

[4]Research Center Trustworthy Data Science and Security, Dortmund, Germany info@sven-mayer.com

We present the *RoboFaceMaker*, an open-source tool for designing and deploying expressive facial animations on screen-based robots. The system allows users to design and animate human-inspired facial features such as eyes, mouth, pupils, and colors, while also supporting symbolic variations such as background changes. The editor also supports transitions between facial expressions and easy deployment across arbitrary screen sizes on physical robots through ROS2. We refined the tool through an iterative process involving three studies: a usability study ($N = 16$), an expert design study ($N = 8$), and a validation study of expression interpretability ($N = 12$). The result is a system that lowers the barrier to authoring robot facial expressions while the created animations can be readily deployed and studied on screen-based robots with ROS2.

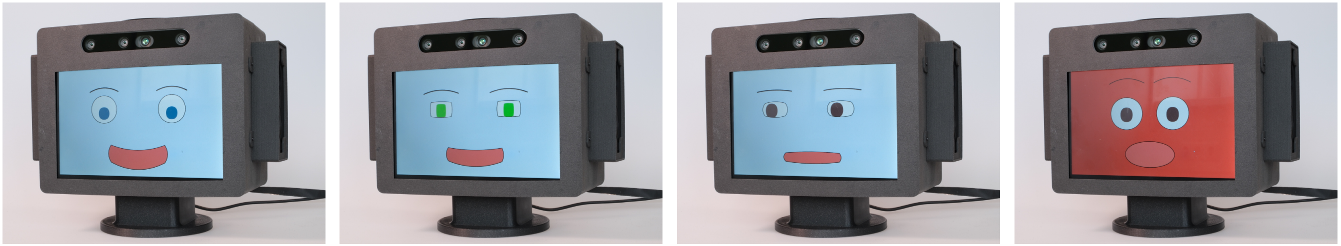
Our evaluations show that the *RoboFaceMaker* supports expert-level authoring, while the validation study revealed an intriguing pattern: expressions that experts found harder to design (as they lacked direct human analogies) were often interpreted more clearly than human-inspired ones. Symbolic cues, such as changes in background color, proved especially effective for communicating intentions. These results highlight both the reliability of our system and its value as a platform for HRI research. Through this, *RoboFaceMaker* enlarges the landscape of existing tools, with a timeline-based animation tool, and real-time ROS2 deployment on robots. By lowering the barrier to designing and deploying facial animations and by releasing the tool and example animations openly, it enables HRI researchers to systematically investigate how robot facial expressions influence trust, coordination, and communication.

II. Related Work

Here, we discuss expressive communication, examine approaches to designing robot facial expressions and their perception, and finally motivate the need for *RoboFaceMaker*.

A. Expressive Communication in HRI

Non-verbal communication is central to human interaction, conveying up to 87% of information in face-to-face exchanges [16]. Facial expressions are powerful for signaling emotions and intentions [5], [6], making them crucial for fostering trust, coordination, and engagement. Robots that lack expressive capabilities risk appearing opaque or unpredictable, reducing user trust and hindering collaboration [11]. Moreover, other non-verbal strategies, such as gaze, body posture, gestures, and acoustic signals, strongly shape how robots are perceived. Direct social gaze, for example, can



(a) Greeting (b) Information (c) Interaction (d) Warning

Fig. 1: Pictures of the four intent-communication facial expression animations on the robotic head used for the validation study.

signal clear intent and increase user comfort [17], while combining facial, gestural, and postural cues enhances emotion recognition and perceptions of trust and naturalness [7], [8], [18]. These effects depend on context and user traits, with well-coordinated, context-aware behaviors leading to more engaging and effective interactions [19], [20].

B. Robot Facial Design

The robot's face design is a critical decision, as it strongly shapes how users perceive and engage with the robot [13], [21], [22]. Faces serve as social cues, influencing trust, perceived competence, and emotional connection [10], [23]. Robot faces can be categorized into several types [21], [22]. Realistic faces mimic human anatomy and enable highly naturalistic expressions but provoke the uncanny valley effect [24]. Symbolic or iconic faces use simplified, abstract features to convey emotions, avoiding uncanny valley effects while remaining approachable and reliable. Other types include blank or mask faces, which suppress identity to allow user interpretation, tech faces, which highlight mechanical components, and screen-based faces, which leverage digital displays to create flexible, dynamic expressions but lack physical embodiment. For screen-based faces, users favor designs with pupils, eyelids, and especially smiling mouths, which are associated with friendliness and sociability [25]. Users who design their own robot faces report greater engagement and affinity for them [26]. Kalegina et al. [13] provide an analysis of 76 robot faces, mapping how specific design elements influence their perception. These findings emphasize that the expressiveness and configurability of facial features are central to designing robots that communicate intentions effectively.

C. Existing Facial Animation Tools

Tools and approaches have been developed for generating robotic facial expressions, ranging from fully automated pipelines to semi-manual design systems. Early work focused on real-time synthesis of facial expressions across different robotic platforms. For instance, an automatic ROS-based approach for facial expression synthesis was presented and evaluated on both a virtual face and a 16-DOF physical robot [27]. Others build pipelines for automatic lip sync or by mapping facial action units to robotic actuators, providing a foundation for consistent expression reproduction [28], [29]. Recent approaches leverage advanced techniques such as LLMs to generate context-aware, dynamic expressions [15], or use vision-based pipelines to map human facial movements directly to robotic heads [30]. These systems primarily focus on translating or generating expressions rather than empowering designers to create, iterate, and refine animations. FLEX SDK [12] and Vizij [31] are the closest to our work, offering open-source platforms that, among other features, support the design and animation of robotic faces.

In summary, existing approaches often rely on automated expression generation or offer limited, pre-defined animations. RoboFaceMaker addresses this by combining iterative, user-driven design workflows with direct, manual animation of facial features with real-time integration.

III. Tool Design and Implementation

We designed and developed the *RoboFaceMaker* to make the creation of expressive robot facial animations more accessible and to easily deploy them on physical robots. Our tool combines a face designer with a ROS2 package for seamless integration into robotic systems. We developed this tool through an iterative process, informed by two user studies and guided by three key goals: **Accessibility:** Users should be able to create animations without coding or robotics expertise. **Rapid Iteration:** Users should be able to prototype, test, and refine facial animations quickly, supporting user-centered HRI research. **Real-World Deployment:** Animations must run on actual robots using open standards and robotics middleware.

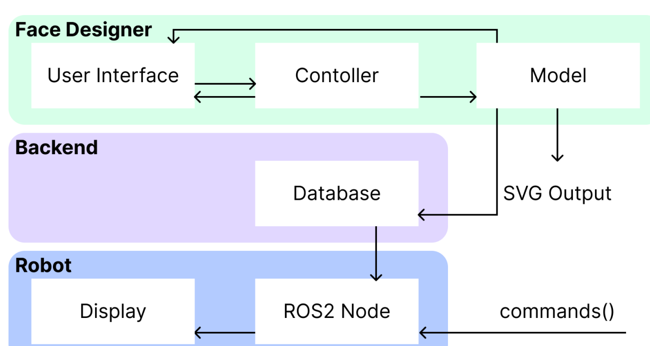


Fig. 2: Overview of the ROS2 Node, with the Architecture of the RoboFaceMaker and User Interaction.

Fig. 3: The RoboFaceMaker web interface to design new faces. a) The player for the designed face, b) the menu options to control the face, and c) the timeline feature to control the keyframes and thus the animations.

A. System Overview

The RoboFaceMaker consists of three main components: (1) **Animation Editor:** A browser application enabling users to design rendered faces and dynamic facial animations using keyframes. (2) **Animation Player**: A browser application in which created and saved animations can be played back. (3) **ROS2 Integration:** A companion package that streams animations to robots in real time through `rosbridge` for immediate testing on physical hardware. Figure 2 illustrates the overall workflow: users create animations in the browser, export them as JSON files, and deploy them directly to a robot with a display-based face. This architecture can be used with any robot that supports ROS2 and has a screen-based face. This supports reproducible HRI experiments by allowing the same animation files to be deployed across labs, platforms, and display sizes without reimplementation.

B. Face Designer

The *Face Designer* provides direct control over facial features (see Figure 3a). The design is inspired by other open-source tools (e.g., [32]). The editor provides a default face to give designers an immediate starting point, lowering the entry barrier. Users can then adjust parameters such as shape, position, and color of facial features. These features align with prior work on robot face design, which highlights eyes, pupils, and mouths as critical for conveying emotions and intentions in interaction [13], [33]. We allow direct, fine-grained control (see Figure 3b) to manipulate cues that strongly influence user perceptions of friendliness, competence, and intent.

The editor includes a timeline view where keyframes can be created and interpolated into smooth animations (see Figure 3c). This enables control over which features change and the temporal dynamics of these changes, a factor known to shape how expressions are interpreted in HRI. The editor also supports saving keyframe presets to the backend database for reuse and storing animations in JSON format to ensure reproducibility and portability across robots and display sizes.

The standalone web interface used only plain HTML and JavaScript. This ensures easy deployment, lowering barriers to adopting the tool across different setups and robots. The resulting face is an SVG vector graphic which avoids rigid bitmap processing and enables precise control and interoperability across devices and screen sizes.

C. Deployment via ROS2 Integration

Animations are stored in the backend database as JSON that describes the positions, shapes, and colors of facial elements over time. This structured format ensures portability and reproducibility: animations can be edited, shared across labs, and reused in different studies. For deployment, we developed a ROS2 package that streams animations in real time via `rosbridge`. The animations are rendered in a kiosk-style web application for display on any robot face screen using ROS2 support. This architecture is robot-agnostic, so the same animation created in the editor can be transferred to robots of different sizes and display configurations without modification.

D. Open-Source Repository

To support reproducibility and adoption within the HRI community, we release the *RoboFaceMaker* as an open-source tool: `https://github.com/mimuc/RoboFaceMaker`

IV. Usability Study

In our within-subjects study, we evaluated the usability of the initial RoboFaceMaker to guide the next iteration of the tool. Our goal was to understand whether users could effectively interact with the interface, whether the provided features were sufficient to create meaningful facial animations, and which challenges arose during the design process. Participants completed four tasks representing four intents. This design allowed us to directly compare performance across contexts while minimizing individual differences. We counterbalanced the task order using a Latin Square design.

A. Task

We based our four tasks on the four intent categories: *State*, *Attention*, *Motion*, and *Instruction* by Pascher et al. [34]. Participants had to create a facial expression for a domestic robot in a household context: a greeting, an informative expression, an expression to initiate interaction, and a warning.

B. Apparatus

We conducted the study in a controlled lab setting. Each participant worked with two monitors: one displayed the RoboFaceMaker interface and task description, while the other displayed surveys and questionnaires. Participants interacted with the tool using a standard mouse and keyboard.

C. Procedure

We began each session by welcoming participants, explaining the study, and obtaining informed consent. Participants then completed a demographic questionnaire and the Affinity for Technology Interaction (ATI) scale [35]. To familiarize them with the interface, we provided a short tutorial followed by a warm-up task in which they created a simple "happy face." In the main experiment, we introduced the robot as a household assistant and presented participants with four tasks. Each task required them to create a new facial expression within max 7 minutes. After each task, they completed an experience questionnaire. Finally, we conducted a semi-structured interview to gather qualitative feedback and suggestions for improving the tool.

D. Measures and Analysis

We combined quantitative metrics with qualitative insights to evaluate the tool. We measured perceived usability using the System Usability Scale (SUS) [36] and cognitive workload using the raw NASA Task Load Index (NASA-TLX) [37]. Through open-ended questions and semi-structured interviews, we explored missing features, usability challenges, and user strategies for designing facial expressions. We also analyzed the generated animations, examining keyframe placement and parameter variations to understand how participants expressed different intentions. We analyzed quantitative data using Python and R, and qualitative data using Atlas.ti.

E. Participants

We recruited 16 participants (8 female, 8 male) aged 19–29 ($M = 23.63$, $SD = 2.70$) from 3 nationalities (German, Bulgarian, Chinese), and education levels, ranging from high school (9), bachelor's degree (5), to master's degrees (2). Five participants had interacted with robotic systems before, and 11 had never interacted with one. Their ATI score[35] was moderate to high ($M = 3.69$, $SD = 0.79$).

F. Results

Participants rated the RoboFaceMaker as highly usable, with an average SUS score of 81.21 ($SD = 12.79$), indicating "excellent" usability [38]. NASA-TLX workload ratings were low ($M = 4.62$, $SD = 2.14$), indicating that participants could interact with the tool without excessive effort. Statistical analysis did not reveal any statistical differences in the task on usability and workload.

We conducted thematic analysis of the open-ended responses and semi-structured interviews, which provides insights into how they experienced the tasks, which features they used or missed, and their impressions of the RoboFaceMaker. We identified three key themes: *Task Evaluation*, *Use and Needs of Features and Elements*, and *Usability*.

Task Evaluation. Participants varied in how they perceived task difficulty. Most found the *Greeting* and *Warning* tasks easiest, as they involved clear, familiar expressions. The Greeting task was described as straightforward because it required creating a simple happy face. P3 explained, "*The easiest one was the greeting task because you just had to make a happy face.*" The Warning task was also intuitive for many, often using red to indicate danger. P1 explained, *The easiest one was the warning task because I could just use red colors to show the warning signals.*" However, others struggled to represent danger without additional context. The *Interaction* task, where the robot had to signal upcoming movement, was considered as the most difficult by half of the participants. They found it hard to convey motion intentions through facial expressions alone. P4 commented, "*The hardest was the Interaction Task because it's something that humans would just talk about and don't use their facial expression for.*" In general, participants found intentions with clear visual cues easier to design, while abstract intentions were harder to express. Using the tool for a while reduced the mentioned challenges.

Usability and Needs of Features and Elements. Most participants described the RoboFaceMaker as intuitive and easy to use. P9 stated, "*It is definitely easy to use. A lot of things are self-explanatory.*" Participants also reported that they learned the system quickly and enjoyed the creative process of animating facial expressions. P3 highlighted this sense of engagement: "*I had a lot of fun doing it with the time frames because you could see the movement live.*" However, a few participants criticized the limited complexity and realism of the facial designs. Across all tasks, participants relied most heavily on the eyes and mouth to create expressions, with the eyes being the dominant feature. Color played a particularly important role in the Warning task, where nearly all participants used red to indicate danger. P2 emphasized the expressive importance of the eyes: "*I would say the eyes, because the eyes are talking as much as they can. The impressions are the most in the eyes.*" Despite these tools, many participants felt limited by the available elements. Eyebrows were the most frequently mentioned missing feature, mentioned by fifteen participants as crucial for conveying emotions. P10 explained, "*The elements I missed most were the eyebrows. I think they would play a crucial part in expressing the faces of the robots.*" Other, less often desired elements included cheeks, noses, and wrinkles. Beyond visual elements, participants requested workflow improvements such as the ability to copy or move keyframes, an undo button, and additional presets to speed up the design process. P3 suggested, "*Maybe you could add a feature that lets you copy and paste the robot faces to a different time frame, because I had to change every keyframe individually.*"

G. Implementation Iteration

Guided by the insights from this study, we made substantial improvements to the RoboFaceMaker. We added adjustable eyebrows to increase emotional clarity, implemented an undo and redo system to support exploration, introduced drag-and-drop timeline editing for precise keyframe management, and enabled users to save and reuse keyframes as presets.

V. Expression Design Study

After implementing the identified improvements, we conducted a second evaluation with domain experts. The second study focused on expressiveness, i.e., whether experts could create high-quality facial animations that effectively conveyed their intended robot intentions. To capture the depth and quality of each expert's design process, we employed a between-subjects design in which each participant designed one of the four intent categories as described in [34].

A. Apparatus

We conducted the study using Zoom. Thus, participants interacted with the improved browser-based tool while sharing their screens. We recorded their screen and the audio to capture their design process and think-aloud comments.

B. Procedure

We began by welcoming the participant, explaining the study's purpose, and obtaining informed consent. Then participants completed a demographic questionnaire, the ATI scale [35], and self-assessed their HRI, HCI, and UX expertise. We then introduced the RoboFaceMaker through a short live demo of key features (sliders, keyframes, timeline editing), and the preset system. We encouraged participants to ask questions and explore the interface before beginning. During the design phase, participants worked independently while thinking aloud, verbalizing their reasoning. Upon task completion, they send us their animations and presets, and completed a post-task survey (SUS [36], NASA-TLX [37], UEQ-S [39], and user satisfaction with self-defined items). We concluded with a semi-structured interview. Sessions lasted 30 to 75 min.

C. Participants

We recruited 8 domain experts (age range 27–58, $M = 35.63$) via our network, who did not participate in the first study. Six participants held doctoral degrees, and two held Master's degrees. Participants came from five different nationalities (German, French, Canadian, Vietnamese, and Colombian). Participants interacted with robotic systems before. ATI scores confirmed that they were highly comfortable with technologies ($M = 5.06$, $SD = 0.54$). All participants had backgrounds and expertise in HCI, HRI, or UX.

D. Quantitative Results

Experts rated usability, rated via the SUS [36], positively ($M = 76.56$, $SD = 11.80$), which corresponds to a "good" usability range [38]. Workload, measured by the raw NASA-TLX [37], was moderate ($M = 6.63$, $SD = 3.70$). User experience, measured via the UEQ-S [39], was rated positive (threshold $>.8$ [39]) for both pragmatic qualities ($M = 1.27$, $SD = 0.49$) and hedonic qualities ($M = 1.17$, $SD = 0.56$). Participants rated their satisfaction high, see Table I.

TABLE I: Results from the eight self-defined satisfaction questions rated with VAS sliders ranging from 0 (strongly disagree) to 100 (strongly agree).

Question	M	SD
I was very satisfied with the animation I created	81.9	5.4
The animation perfectly matched my imagination.	75.9	18.4
The animation I created is very suitable for the intended purpose.	89.9	10.1
The quality of the final animation is very high	65.6	23.1
The animation creation process met my expectations very well.	81.3	13.3
I feel like I completed the task very successfully	88.3	12.3
I was very successful in creating the animation I intended	84.9	14.3
The process of creating the animation felt very intuitive to me	70.8	31.6

E. Analysis of Created Animations.

We analyzed the final animations to better understand how experts visually expressed the four robot intentions. On average, participants used 4.88 keyframes ($SD = 2.80$). All participants used the mouth, eyes, pupils, and eyebrows in all keyframes. Participants mainly used the default background color (bright yellow/white) for the greeting, interaction, and information intention, except for one participant who used a bright yellow for the interaction expression. For the warning expression, both participants used a gradient from the default background color to red. Across all animations, eyes were mostly round. For all keyframes, we found that the mouth and eyes were the most frequently manipulated features, with eyebrows and pupils used more selectively, see Figure 4.

Greeting Intention: Both participants used a smiling mouth that opened and closed, resembling a verbal "hello." Eyes were round and bright, with thin, slightly curved eyebrows, and movements were slow and continuous. One expert added a wink of one eye after the mouth movement. **Interaction Intention:** This involved faster, more explicit movements of both the eyes, pupils, and mouth. The independent pupil movement suggested the robot first looked around before "speaking." Eyes were wider horizontally, and animations had multiple stages. **Warning Intention:** This animation was visually distinct, using a background gradient shifting to deep red, wide-open mouths, and color shifting and moving eyes. The eyebrows moved closer to the eyes, with all features moving in sync. **Informative Intention:** This expression showed the highest variation. One animation used smooth, continuous movements of the mouth and eyebrows to express curiosity, while the other relied on discrete, stepwise changes in the eyes to indicate

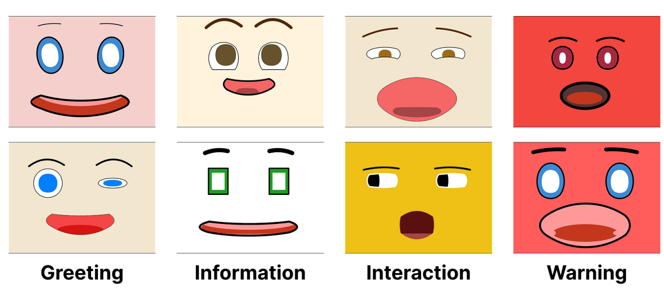


Fig. 4: Screenshots of resulting facial expressions.

different informational states. In both cases, the eyes and mouth moved independently.

Based on this, two authors collaboratively used open coding to combine the two expressions for each intention into one, taking into account the main features from both experts [40]. The four resulting expressions had a consistent style to reduce confounds and confusion (see Figure 1). We then continued the validation study with these merged expressions.

VI. Expression Validation Study

To confirm whether the expressions created by the experts are actually understandable to others, we adopted the expression elicitation approach [40]. This approach suggests first collecting expressions based on referents, combining these expressions into their core ideas, and then, in a second study, validating whether a different set of participants understands the expressions in line with the referent they were created for. We used a within-subject design; thus, each participant saw and interpreted the meaning of all four expressions. We counterbalanced the order in which participants saw the animations using a Balanced Latin Square Design and conducted the study as an in-person lab study.

A. Apparatus

We used our ROS2 package to bring these expressions onto a robotic head. For this, we used an open source display-based robotic head [41], see Figure 1. Participants observed the robotic expressions on the robotic head and then filled out questionnaires on a separate PC.

B. Procedure

We began each session by welcoming participants, explaining the study's purpose, and collecting their written informed consent. After this, participants filled out a demographic survey. We then asked participants to imagine that they are in their house with their domestic robot. The robot shows its intent via facial expressions. Participants could then observe one of the facial expressions as often as they wanted to. Afterward, we asked them to fill out a short questionnaire in which they had to interpret what they thought the robot wanted to express and rate the expression on a scale of how understandable, noticeable, disturbing, effective, and intrusive it was. After they were done with all four expressions, we then showed them the intended meanings of the expressions and let them order each of the four intentions to one of the facial animations. We then thanked participants for their participation. Each session lasted approximately 10 minutes.

C. Participants

We recruited 12 (5 male, 6 female, 1 non-binary) participants, who had not taken part in the previous studies, via convenience sampling. Participants' ages ranged from 26 to 33 ($M = 29.67$, $SD = 2.50$). 11 participants had a master's degree, and 1 had a bachelor's degree. Participants' ATI score was 5.02 ($SD = 0.58$). One participant had never interacted, 5 had interacted between 1 and 7 times, and 6 had interacted more than 7 times with robotic systems before. Those who have interacted with robots before, mostly named robotic arms, Roombas, and social robots.

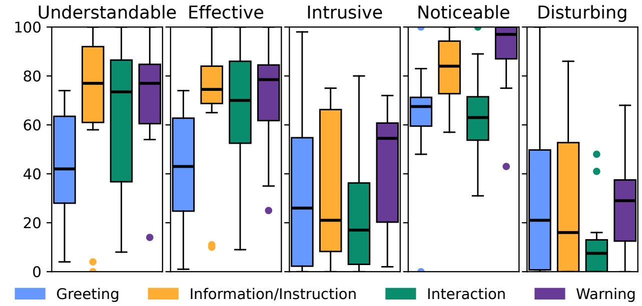


Fig. 5: Perceptions of the four facial expressions. Higher scores depict higher perceived properties.

D. Results

In line with the expression elicitation approach, we coded participants' interpretations into code-groups and calculated the qualitative response accuracy (QRA) [40].

Our iterative coding resulted in 28 labels from 48 expressions. We found 13 code groups (*Attention, Communicate Movement, Communicate Success, Confirmation, Error, Excitement, Expresses Uncomfortable, Friendly, Greeting, Negative Emotion, Request Help/Interaction, Uncertainty, Warning About Danger*). For each of the four intentions, we determined which code groups correspond to the original meaning of the referent used to create the facial expression. With this, we calculated the QRA for each intention, indicating how many participants correctly understood each intention (Greeting: 0.25, Information: 0.67, Interaction: 0.42, Warning: 0.67).

Furthermore, after participants interpreted all four intentions, we presented them with the originally intended intentions and then asked them to match each referent to the four faces. Here, we found that all participants correctly matched the *warning* intent, 11 of 12 correctly matched the *greeting and interaction* intent, and 12 correctly matched the *information* intent. This indicates that, without context, only the *Information* and *Warning* intentions were reliably understood, whereas with prior knowledge, all could be reliably matched.

We asked participants five questions about their perceptions of the facial expressions to gain further insights; see Figure 5. Here we observed that they perceived the *greeting* expression the least understandable and effective, while the other three intentions were rated high. Intrusion and disturbance were rated generally low, with *warning* being perceived as higher. For noticability all were rated high, with *information* and *warning* receiving the highest ratings.

VII. Discussion

In the following, we discuss challenges and implications in communicating intentions via facial expressions and how our tool can support researchers in rapidly designing animations for robot faces.

A. RoboFaceMaker Supports Rapid and Systematic Animation Design

The RoboFaceMaker primarily addresses the issue that designing expressive robot faces is often time-consuming and requires coding skills or predefined templates [12], [15], limiting rapid, iterative, and user-centered design. Our work demonstrates how researchers can design, test, and refine animations in real time, and deploy them directly to robots via ROS2, allowing them to focus on high-level behavior rather than low-level implementation details. Compared to automated generation approaches [15], the tool preserves full authorial control, while being compatible with agentic systems that can trigger expressions at the right time. This combination accelerates early-stage design while maintaining precision in expression creation.

B. Human-Like Expressions vs. Symbolic Designs

Our studies revealed a trade-off between the ease of designing expressions and their reliability of understanding. Human-like expressions, such as smiles or winks, were intuitive for experts but ambiguous for users. In contrast, expressions without human analogs, such as warnings using the background color, were challenging to design but were understood in the validation study. Symbolic cues proved effective for intention communication.

Prior work highlights non-verbal communication for trust and coordination [7], [13]. Our findings extend this by providing initial evidence that symbolic designs can offer more reliable affective grounding [42]. Rather than replicating human affect, grounding expressions in shared conventions (for example, red for danger or green for readiness, i.e., traffic light metaphor) creates a representational language that reduces ambiguity and supports clearer communication of robot intentions. At the same time, such conventions are not universal and may vary across different cultural contexts, which future research should systematically explore.

Human-like expressions remain valuable in domains where familiarity, empathy, and social presence are essential, such as care, education, or companionship. However, simplified robot expressions cannot capture the nuance of human affect and may lead to misinterpretation or even uncanny effects if expectations are not met [24], [21]. Symbolic strategies, in contrast, are more discrete and less context-dependent, making them suitable in settings where safety and transparency are critical. These categories are not mutually exclusive, though; hybrid designs that combine symbolic cues with human-inspired features are a promising middle ground.

Our empirical results highlight the importance of context. Symbolic strategies appear more reliable for safety-oriented scenarios, while human-like designs remain engaging in social environments.

C. Enabling a New Research Paradigm

Recent work in HRI has begun to emphasize systematic approaches for designing and validating robot expressions. Expression elicitation methods (e.g., Leusmann et al. [40]) and related methods in non-verbal behavior design [43], [44] highlight the importance of reproducible, human-understandable expressions. These approaches shift the focus from ad hoc or emotionally inspired designs toward standardized, testable procedures.

Therefore, we provide an open tool chain for authoring, deploying, and validating screen-based robot expressions. *RoboFaceMaker* enables researchers to explore diverse expression strategies, human-like, symbolic, or hybrids, and to compare them systematically across studies and platforms. This creates new opportunities for cumulative knowledge building in HRI, moving the field toward more replicable and theory-driven research on robot expression design.

D. Limitations and Future Work

Our work has several limitations that should be considered when interpreting the findings. First, the studies involved small sample sizes: eight experts designing faces and 12 participants in the validation study. These findings should be seen more as preliminary empirical findings and further validated. Second, we focused on four intentions [34], to ensure theoretical grounding. While this allowed for focused exploration, a broader set of intentions is needed to fully understand how facial expressions can communicate a wide range of robot states and intentions. Third, more user-centered design iterations could reveal additional needs and refinements. For example, we plan to support pre-configured animations, such as blinking. Finally, although we designed the system for screen-based robots, it can also be deployed in AR contexts through Unity streaming, enabling future exploration of AR faces [33], [45], [46].

VIII. Conclusion

We introduced RoboFaceMaker, an open-source tool that enables researchers to design and deploy expressive facial animations for screen-based robots through a browser-based editor with ROS2 integration. We show that the system supports users in creating expressive facial animations for physical robots. The findings provide evidence that symbolic designs, such as color changes, may sometimes be clearer for conveying intention than human-like expressions, although further work is needed to explore their use across different contexts. By lowering the prototyping barrier and supporting reproducible experimentation, RoboFaceMaker offers a foundation for systematic research on how robot faces can foster trust, coordination, and communication in HRI.

References

[1] S. Panagou, W. P. Neumann, and F. Fruggiero, "A scoping review of human robot interaction research towards industry 5.0 human-centric workplaces," *J. of Production Research*, 2024. https://doi.org/10.1080/00207543.2023.2172473

[2] M. Kyrarini, F. Lygerakis, A. Rajavenkatanarayanan, C. Sevastopoulos, H. R. Nambiappan, K. K. Chaitanya, A. R. Babu, J. Mathew, and F. Makedon, "A survey of robots in healthcare," *Technologies*, 2021. http://doi.org/10.3390/technologies9010008

[3] I. Lee, "Service robots: A systematic literature review," *Electronics*, 2021. https://doi.org/10.3390/electronics10212658

[4] C. Mavrogiannis, F. Baldini, A. Wang, D. Zhao, P. Trautman, A. Steinfeld, and J. Oh, "Core challenges of social robot navigation: A survey," *J. HRI*, 2023. https://doi.org/10.1145/3583741
[5] P. Ekman and W. V. Friesen, "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding," *Semiotica*, 1969. https://doi.org/10.1515/semi.1969.1.1.49
[6] I. M. Menne and F. Schwab, "Faces of Emotion: Investigating Emotional Facial Expressions Towards a Robot," *J. of Social Robotics*, 2018. https://doi.org/10.1007/s12369-017-0447-2
[7] C. Tsiourti, A. Weiss, K. Wac, and M. Vincze, "Designing Emotionally Expressive Robots: A Comparative Study on the Perception of Communication Modalities," in *I. C. on Human Agent Interaction*. ACM, 2017. https://doi.org/10.1145/3125739.3125744
[8] M. Salem, K. Rohlfing, S. Kopp, and F. Joublin, "A friendly gesture: Investigating the effect of multimodal robot behavior in hri," in *RO-MAN*. IEEE, 2011. https://doi.org/10.1109/ROMAN.2011.6005285s
[9] H. Lehmann, A. V. Sureshbabu, A. Parmiggiani, and G. Metta, "Head and face design for a new humanoid service robot," in *Social Robotics*. Springer, 2016. https://doi.org/10.1007/978-3-319-47437-3_37
[10] Y. Song, A. Luximon, and Y. Luximon, "The effect of facial features on facial anthropomorphic trustworthiness in social robots," *Applied Ergonomics*, 2021. https://doi.org/10.1016/j.apergo.2021.103420
[11] H. Ayanoğlu and E. Duarte, Eds., *Emotional Design in Human-Robot Interaction: Theory, Methods and Applications*, ser. HCI Series. Springer, 2019. http://doi.org/10.1007/978-3-319-96722-6
[12] P. Alves-Oliveira, K. Mihata, R. Karim, E. A. Bjorling, and M. Cakmak, "FLEX-SDK: An Open-Source Software Development Kit for Creating Social Robots," in *UIST*. ACM, 2022. https://doi.org/10.1145/3526113.3545707
[13] A. Kalegina, G. Schroeder, A. Allchin, K. Berlin, and M. Cakmak, "Characterizing the design space of rendered robot faces," in *HRI*. ACM, 2018. https://doi.org/10.1145/3171221.3171286
[14] G. J. Serfaty, V. O. Barnard, and J. P. Salisbury, "Generative facial expressions and eye gaze behavior from prompts for multi-human-robot interaction," in *Adjunct UIST*. ACM, 2023. https://doi.org/10.1145/3586182.3616623
[15] V. N. Antony, M. Stiber, and C.-M. Huang, "Xpress: A system for dynamic, context-aware robot facial expressions using language models," in *HRI*. IEEE, 2025. https://doi.org/10.1109/HRI61500.2025.10974045
[16] J. K. Burgoon, V. Manusov, and L. K. Guerrero, *Nonverbal Communication*. Routledge, 2016. https://doi.org/10.4324/9781315663425
[17] F. Vigni and S. Rossi, "Exploring Non-verbal Strategies for Initiating an HRI," in *Social Robotics*. Springer, 2022. https://doi.org/10.1007/978-3-031-24667-8_25
[18] X. Zhu, X. Zhang, Z. Chen, Z. Dong, Z. Gu, and D. Chang, "The Trusted Listener: The Influence of Anthropomorphic Eye Design of Social Robots on User's Perception of Trustworthiness," in *CHI*. ACM, 2022. https://doi.org/10.1145/3491102.3517670
[19] P. Aliasghari, M. Ghafurian, C. L. Nehaniv, and K. Dautenhahn, "Impact of nonverbal robot behaviour on human teachers' perceptions of a learner robot," *Interaction Studies*, 2021. https://doi.org/10.1075/is.20036.ali
[20] C. L. Bethel and R. R. Murphy, "Non-Facial and Non-Verbal Affective Expression for Appearance-Constrained Robots Used in Victim Management*," *Paladyn, Journal of Behavioral Robotics*, 2010. https://doi.org/10.2478/s13230-011-0009-5
[21] M. Blow, K. Dautenhahn, A. Appleby, C. L. Nehaniv, and D. Lee, "The art of designing robot faces: dimensions for hri," in *C. on HRI*. ACM, 2006. https://doi.org/10.1145/1121241.1121301
[22] C. Chesher and F. Andreallo, "Robotic faciality: The philosophy, science and art of robot faces," *J. of Social Robotics*, 2021. https://doi.org/10.1007/s12369-020-00623-2
[23] D. Peña and F. Tanaka, "Human Perception of Social Robot's Emotional States via Facial and Thermal Expressions," *ACM Transactions on HRI*, 2020. https://doi.org/10.1145/3388469
[24] M. Mori, K. F. MacDorman, and N. Kageki, "The Uncanny Valley [From the Field]," *Robotics & Automation Magazine*, 2012. https://doi.org/10.1109/MRA.2012.2192811
[25] M. Ringwald, P. Theben, K. Gerlinger, A. Hedrich, and B. Klein, "How Should Your Assistive Robot Look Like? A Scoping Review on Embodiment for Assistive Robots," *Journal of Intelligent & Robotic Systems*, 2023. https://doi.org/10.1007/s10846-022-01781-3
[26] T. Heuer, "Who Do You Want to Talk to?: User-centered Design for Human-like Robot Faces," in *Mensch Und Computer 2019*. ACM, 2019. https://doi.org/10.1145/3340764.3344878
[27] M. Moosaei, C. J. Hayes, and L. D. Riek, "Performing facial expression synthesis on robot faces : A real-time software system," 2015. https://api.semanticscholar.org/CorpusID:14339761
[28] R. Amini, C. Lisetti, and G. Ruiz, "HapFACS 3.0: FACS-Based Facial Expression Generator for 3D Speaking Virtual Characters," *Transactions on Affective Computing*, 2015. https://doi.org/10.1109/TAFFC.2015.2432794
[29] N. S. Dennler, E. Torrence, U. Yoo, S. Nikolaidis, and M. Mataric, "Pylips: An open-source python package to expand participation in embodied interaction," in *S. on User Interface Software and Technology*. ACM, 2024. https://doi.org/10.1145/3672539.3686747
[30] M. Heisler and C. Becker-Asano, "An iPhone pro is all you need: Mimicking facial expressions on an android robot head," in *HRI*. IEEE, 2025.
[31] S. Elbeleidy, A. Schoen, C. Birmingham, T. Ribeiro, V. Paléologue, D. Dooley, and R. Mead, "Vizij: A workshop & tutorial on expressive robot faces," in *Companion HRI*. ACM, 2026. https://doi.org/10.1145/3776734.3789491
[32] V. Schwind, K. Wolf, and N. Henze, *FaceMaker—A Procedural Face Generator to Foster Character Design Research*. Springer, 2017. https://doi.org/10.1007/978-3-319-53088-8_6
[33] J. E. Young, M. Xin, and E. Sharlin, "Robot expressionism through cartooning," in *HRI*. ACM, 2007. https://doi.org/10.1145/1228716.1228758
[34] M. Pascher, U. Gruenefeld, S. Schneegass, and J. Gerken, "How to communicate robot motion intent: A scoping review," in *CHI*. ACM, 2023. https://doi.org/10.1145/3544548.3580857
[35] C. A. Thomas Franke and D. Wessel, "A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale," *J. of HCI*, 2019. https://doi.org/10.1080/10447318.2018.1456150
[36] J. Brooke, "SUS-A quick and dirty usability scale," *Usability evaluation in industry*, 1996.
[37] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later," *Human Factors and Ergonomics Society Annual Meeting*, 2006. https://doi.org/10.1177/154193120605000909
[38] J. Brooke, "Sus: a retrospective," *Journal of Usability Studies*, 2013.
[39] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S)," *J. of Inter. Multimedia and AI*, 2017. https://doi.org/10.9781/ijimai.2017.09.001
[40] J. Leusmann, S. Villa, T. Liang, C. Wang, A. Schmidt, and S. Mayer, "An approach to elicit human-understandable robot expressions to support hri," in *CHI*. ACM, 2025. https://doi.org/10.1145/3706598.3713085
[41] J. Leusmann, J. Wiese, M. Ziarko, and S. Mayer, "Investigating opportunities for active smart assistants to initiate interactions with users," in *I. C. on Mobile and Ubiquitous Multimedia*. ACM, 2023. https://doi.org/10.1145/3626705.3631787
[42] M. F. Jung, "Affective grounding in hri," in *HRI*. ACM, 2017. https://doi.org/10.1145/2909824.3020224
[43] A. B. V. D. Nguyen, J. Leusmann, S. Mayer, and A. Vande Moere, "Eliciting understandable architectonic gestures for robotic furniture through co-design improvisation," in *HRI*. IEEE, 2025. https://doi.org/10.1109/HRI61500.2025.10973858
[44] G. Venture and D. Kulić, "Robot expressive motions: A survey of generation and evaluation methods," *J. HRI*, 2019. https://doi.org/10.1145/3344286
[45] Y. Hiroi, S. Hisano, and A. Ito, "Evaluation of head size of an interactive robot using an augmented reality," in *World Automation Congress*. IEEE, 2010.
[46] C. Wang, M. Gienger, and F. Zhang, "Generation of real-time robotic emotional expressions learning from human demonstration in mixed reality," 2025. https://arxiv.org/abs/2508.08999