ACM Transactions of Human Robot Interaction, 2021

Integrating Robot Manufacturer Perspectives into Legible Factory Robot Light Communications

ALEXANDRA BACULA, Collaborative Robotics and Intelligent Systems Institute, United States
JASON MERCER, OTTO Motors Autonomy Team, Canada
JADEN BERGER, Collaborative Robotics and Intelligent Systems Institute, United States
JULIE A. ADAMS, Collaborative Robotics and Intelligent Systems Institute, United States
HEATHER KNIGHT, Collaborative Robotics and Intelligent Systems Institute, United States

Social robotics researchers worked with a robotics company consisting of 250 employees whose robots support point-to-point transport in factories and warehouses. Unlike other industry models, these robots are designed to operate with people in mixed human-machine spaces, yet have not been explicitly designed with human-robot communication in mind. Thus, the goal of this work was to elicit input for improving this robot's communication and decision-making strategies around people, and apply this input to improve robot legibility. To achieve this, (1) a company-wide survey relative to the robot's light, sound, and motion communications was sent out and analyzed , (2) three new light sets were developed integrating employee input as well as social robotics research results, (3) these light sets were evaluated relative to the current robot default light patterns, all significantly improving the overall legibility of robot state: at goal, blocked, turning, idle, (4) finally, the latest software release for this robot has deployed a subset of these light patterns to all of their currently operating client sites, i.e., anyone who updates their robots to the latest release will benefit from these research results.

CCS Concepts: • **Human-centered computing** → *Interaction techniques*; *Participatory design.*

Additional Key Words and Phrases: datasets, neural networks, gaze detection, text tagging

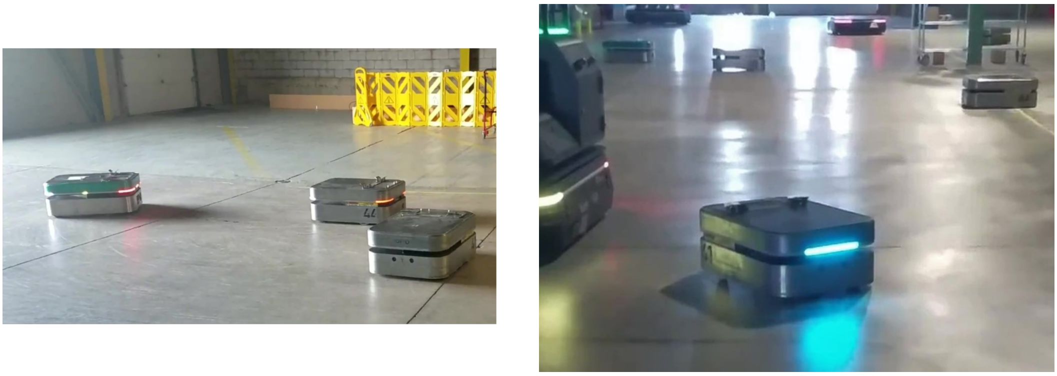
ACM Reference Format:
Alexandra Bacula, Jason Mercer, Jaden Berger, Julie A. Adams, and Heather Knight. 2020. Integrating Robot Manufacturer Perspectives into Legible Factory Robot Light Communications. *ACM Trans. Hum.-Robot Interact.* X, X, Article X (X 2020), 32 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Clearpath OTTO robots operate successfully on factory floors, side-by-side with human workers. The OTTO robots move autonomously without barriers or distinct robot-only zones, navigating around humans and completing tasks while still keeping the humans safe. While many factories utilize mobile robots, such as the Kiva robots at Amazon [24], the robots often operate in a space

Authors' addresses: Alexandra Bacula, baculaa@oregonstate.edu, Collaborative Robotics and Intelligent Systems Institute, 1680 SW Monroe Ave, Corvallis, Oregon, United States, 97331; Jason Mercer, OTTO Motors Autonomy Team, 1425 Strasburg Rd, Kitchener, Ontario, Canada, jmercer@clearpath.ai; Jaden Berger, Collaborative Robotics and Intelligent Systems Institute, 1680 SW Monroe Ave, Corvallis, Oregon, United States, 97331, bergejad@oregonstate.edu; Julie A. Adams, Collaborative Robotics and Intelligent Systems Institute, 1680 SW Monroe Ave, Corvallis, Oregon, United States, 97331, julie.a.adams@oregonstate.edu; Heather Knight, Collaborative Robotics and Intelligent Systems Institute, 1680 SW Monroe Ave, Corvallis, Oregon, United States, 97331, knighth@oregonstate.edu.

2573-9522/2020/X-ARTX $15.00
https://doi.org/10.1145/1122445.1122456



(a) The Clearpath warehouse with OTTO 100s

(b) Image of multiple varieties of OTTO robots (100 and 1500) utilizing their lights during operation.

Fig. 1. Examples of OTTO robots in the Clearpath Warehouse

separate from the humans, or travel on designated paths and do not stray [27]. Thus, the OTTO robots offer unique benefits in that they can be directly integrated into existing factory and warehouse spaces. In doing this, however, it is imperative that such robots are not only able to safely operate around people, but also to operate in a way that people can understand so that they are able to work efficiently in the presence of factory robots. For example, if a person is standing in the way of a robot on the factory floor, the robot needs to be able to clearly communicate a message that tells the person to move.

The goal of this paper is to develop principled robot communications that (1) cater to the explicit task domain of object transport in factory settings, and (2) leverage previous knowledge, that is, to integrate insights from the robot manufacturers and prior work in social robotics. Our previous work with Clearpath found that OTTO robot operations are judged relative to what a person would be doing in a similar situation [10]. Similarly to what Clearpath has observed, social robotics researchers have also considered the benefits of communicative, predictable robots. It has been shown, for example, that people's mental model of what a robot's current task intention is can be influenced by varying the robot's motion en route to a goal Knight et al. [22]. A related study illustrated how robot velocity influences people's willingness to interact with a robot [23]. Further, researchers have developed the concept of "legibiity" to represent people's ability to understand the current state and goal of the robot operations both in motion [8, 12, 25] and in lights [4, 5, 42]. However, these communications could be through not only the robot's movement and light, but also through sound or other channels of communication that humans are able to pick up.

While prior research has underscored the importance of legibility to effective human robot collaboration [11, 17], research results for multi-modal robot communication have seldom included the perspectives of those developing robots for factory contexts. This use of real, deployed factory robots and the insights of the employees of the company that makes them is a novel compare to prior research in designing communicatory signals in robots. Thus, before deciding what communication channel to focus on in our joint work together, the research team began with a survey sent to the company's 250 employees in the summer of 2019. The Clearpath employees offer unique perspectives because of the wide variety of human-machine settings in which the company deploys robots. To capture this perspective, the company survey included a range of scenarios for which they could suggest light, sound, motion or multi-modal robot actions or behaviors. The scenarios were sourced from interactive discussions with the Autonomy and leadership teams at Clearpath, conducted over a year-long period, which included a three month on-site collaboration at the company headquarters, as seen in 1a. The results of this survey provided several ideas for updating the situations and strategies in which the robot communications could be improved, **particularly highlighting the importance of flexibility of factory robot light communications**.

The next step was to develop and implement new expressive light sets based on the results from the company survey alongside previous work in social robotics. Based on a grounded coding analysis of the company survey results, we identified two major groupings of light communication styles suggested by the employees: car-like and sweeping lights. We further extended the situations in which robots would seek to trigger light communications to include idle, blocked, and at goal, whereas previous versions only included turn or drive. We also identified a third light communication strategy from the literature involving glowing light that had shown effective results for when the robot needed help (blue) or was blocked (red) [3, 4] [5]. These three light sets were custom developed for the OTTO robots, and integrated into the existing software infrastructure at the company during the collaboration at Clearpath Robotics. We then conducted a within-subjects user study (N=30) in which employees observed and commented on real robot operations to evaluate the merit of the three custom developed light sets relative to what the company had already been using. These employees had diverse technical backgrounds and familiarity with the robots, mimicking the range of what would be seen in end users.

In comparing the default factory lights to the newly implemented light communications, we found that people valued a wider set of robot state communications. For example, communicating states such as idle, blocked, and task complete was perceived to be helpful in enabling the robot's human coworkers to understand the state of the robot without being overly distracted from the work they needed to do themselves. All three custom light sets outperformed the default light set based on survey responses. It also turned out that **different approaches to light communication fit best to particular state communications**, e.g., less distracting patterns were better for ambient communications without deadlines (like the robot's task being done), whereas noticeable patterns like targeted lights were better if the robot needed something right away. After seeing these results, Clearpath elected to include aspects of the custom light sets into their October 2019 software update. Thus, versions of these new light patterns have been running for tens of thousands of hours across hundreds of robots in the year since these data were collected. Future work will further leverage the collected data to integrate employee perspectives on factory robot sound communications. This work serves as an initial look into the utility of employee insights in designing effective communicative signals for factory robots, which future researches can build upon for their own specific applications.

2 RELATED WORK

Extensive work has been done in social robotics research to improve communicatory abilities for robots in human spaces. This section provides an introduction to previous work in social robotics communication, HRI in factory robots, and participatory design.

2.1 Social Robotics and the Power of Human-Robot Communications

While early social robotics research often focused on robots with anthropomorphic forms, these days, low-DOF robots have a variety of communications channel options, including lights [4, 26, 42], sound [26, 40], and motion [22, 23, 28]. As the present paper focuses on factory robots with no verbal capabilities, these three channels are particularly relevant and were included in our customer survey. Each can be used individually or in tandem to express a variety of communications or robot internal states [26]. For example, even minimal, solid or blinking colored lights can make robot internal states more intuitive to a user [3–5] [16, 41]. This use of lights (or other channels) to clarify the current operational state or goals of the robot to humans is called *legibility* in the literature [12, 25], and was the motivation for expanding the OTTO robot communication capabilities to something that a human co-worker could interpret at a distance. Like light, sound can also be used to indicate state, express emotion, and communicate information Song and Yamada [40] (such as a car beeping to remind you to move forward after the light changes) with prior work considering both functional and additive sound [30, 31]. The final channel of non-verbal communication open to simple mobile robots is motion. Previous work has used motion to demonstrate a robot's attitude toward time (e.g., rushed, relaxed) [21], clarify to whom a robot is attempting to deliver an object [20, 33], or even influence human behavior and interaction patterns around robots [1, 15, 38, 43]. The tradeoff is that robot motion needs to fit into functional robot priorities. Finally, in future work multi-modal expressions can potentially be used in tandem to disambiguate robot communications Löffler et al. [26], providing further fodder for future investigations.

2.2 Prior Studies in Factory Robot HRI

Previous investigations of human-robot interaction in factories reveal important success factors and communication opportunities in factory domains. Much work has been done on how robots can be designed and controlled to allow better collaboration with humans [2, 19, 29, 34, 44], but here we focus particularly on what factors are important to humans for improving that collaboration. One survey by Wurhofer et al. [45] examined many different studies in user design for factory robots to determine what user design factors are most relevant. Of these, perceived safety, trust, and emotions and feelings are highly relevant to the work presented in this paper. Other works have also found that trust is an important factor in robots in industrial settings [18]. In factories that have robots on the floor with humans, studies have shown that communicating intention clearly is important to successful collaboration [6]. In work by Bauer et al. [6], the focus was on how the robot can communicate its intentions to a human, while work by Chun and Knight [10] sent an anthropologist to Clearpath to study how people who work with robots anthropomorphize them.

2.3 Participatory Design

Design is a key part of HRI research, and knowing what users want in a robot they will be interacting with is crucial. Useful methods ranging from participatory design to ethnography. To integrate employee opinions and ideas into our robotic communication, we look to participatory design methods, particularly in robotics, to inform the company survey. Participatory design is a design process that involves collaboration and input from both designers and the intended user group to create designs for products that take into account the wants and needs of users [32, 39]. Participatory design methods can range from ethnographic methods to specific customization of an individual system [32]. In robotics, participatory design has been used in several application areas, with a wide range of participatory design methods being applied successfully to help answer research questions. Fallatah et al. [14] used ethnographic methods to explore how different microcultures across a university campus responded to a robot asking for help, which in turn informed later design decisions. Rose and Björling [36] took a less observatory approach and interacted directly with users in sessions to gain input on designing robots for measuring stress in teens. Šabanović et al. [37] conducted interviews and participatory design workshops for socially assistive robots for elder care with adults who have depression and physical illness. Outside of robotics, Osswald et al. [35] has used creative cultural probing techniques [7] to examine how workers related to the semiconductor manufacturing technology they used, and how they might respond to integrating collaborative user interface technology.

Fig. 2. The OTTO 100 and channels of communication

3 RESEARCH SITE AND TIMELINE

To understand the platform and context in which the light sets were created, this section provides a background on the OTTO robots, manufactured by Clearpath, where the research was conducted for the present study.

OTTO robots are presently used for mobile transport in over 100 factories and warehouses throughout North America, Europe and Asia [10]. There are currently three types of OTTOs: the OTTO 100s, the OTTO 750s, and the OTTO 1500s. The OTTO 1500s and OTTO 750s are large, flat self-driving vehicles that can carry up to 1500kg and 750kg respectively. The OTTO 100 is a small, rectangular box shaped self-driving vehicle that can carry up to 100kg. This work focuses on the OTTO 100s, as the most commonly deployed platform with smaller size, increased agility, and better availability. In the past, employees working closely with these particular robots expressed a strong desire for improved robot communications, which is the dominant motivation for this work [10].

Clearpath currently has around 250 employees, many of whom can be described as either technology developers or customer liaisons. As such, they are well versed in both the capabilities and current communication strategies of the robots, as well as a broad range of customer needs. This knowledge of the contexts in which the robots will be deployed gives the employees broad insights into robot capabilities and user needs, as outlined in Table 1. For example, within client services, the field services team goes to many different end user sites to debug robots, and interacts with the robots and end users in the factories in which the robots are deployed. Within marketing, customer relations regularly speaks to the end users and knows about their grievances and their praise. On the other end of the spectrum, the testing and autonomy teams interact with the robots on the floor daily in tasks that are meant to mimic the tasks the robots perform at the end user sites. All teams also share information with each other, as colleagues and via regular company-wide meetings, also offering collective understandings of user needs and technological possibility.

This broad range of employee roles at Clearpath leads mirrors the large spectrum of familiarity with the robots that factory and warehouse end users and coworkers will have. For example, some employees work on the floor with the robots and interact with the robots every day, which gives them knowledge of how the robots currently work and what they would like to be different. Other employees might rarely interact with the robots, thus it is important to understand the perspective of how clear the robot communications are for a user who only sees the communications at a

TEAM	SUB TEAMS	ROLES	UNIQUE INSIGHTS
Software	Autonomy, Operating Systems, Systems Engineering, Fleet Manager	Design, maintain, and integrate the software to control, manage, task the robots, and allow them to move autonomously	Deep understanding of how the robots navigate through space, perform tasks, and the limitations of the software
Hardware	Mechanical Design, Electrical Design	Design and engineer the mechanical and electrical systems of the robots	Knowledge of the physical capabilities of the robots, and their physical design strengths and limitations
Testing	Testing	Test the capabilities of the robots in a factory floor setup in a range of tasks	First hand experience interacting with and observing the robots perform tasks relevant to end users
Product Design	Product Engineering, Product Applications, Application Engineering, User Experience	Designing new features, products, and applications for the company, design and maintain systems users interact with	Knowledge of product features, contexts in which the robots are used, and how people interact with the robots
Client Services	Field Services, Client Success, Professional Services	Interact with end users to debug, update, troubleshoot, and provide support for the robots	First hand experience in end user sites, direct interaction with end users and robots together
Marketing	Sales, In-Market Engineering, Commercial Management	Selling products to potential end users, marketing the robots in an effective way	Knowledge of what features are most attractive to potential end users, and what current products are selling best

Table 1. Employee teams at Clearpath, their roles, and unique insights from these roles.

low rate. The range of technical expertise in our employee pool is also similar, offering a valuable participant set more akin to industry than an academic setting. Future work can also consider how variants in uses of these robots, like giant factories vs small printing operation, could further inform customized communications, but as a first step, we integrate employees are familiar with this broad range of applications in which the robots are used.

The work presented in this paper occurred from June 2019 to September 2019, with the company survey conducted in July, the technology developed in August, and the user study run in September.

4 THE COMPANY SURVEY

To integrate employee insights into robot expressive communication strategies, we developed a company-wide survey. The goal was to extract needs and ideas for robot expressive communications from the people who best understand the broad range of contexts and the end users that utilize the robots. Why survey the Clearpath employees? As introduced in the previous section, the surveyed employees have a wide spectrum of specialities, from customer support, to robot testing, to building the robots, which may give them unique insights as to where it might be valuable to improve the robots. We hypothesized that surveying across these different groups would allow for a more comprehensive look into what can be improved in the robots' communications. Firstly, as mentioned in Section 3, each team has unique insights into how to improve specific aspects of the robots, which may give more nuanced feedback and a broader range of specific suggestions. Prior work has shown that different teams at Clearpath anthropomorphize the robots differently, which could also contribute to differing opinions [10]. Secondly, the different teams may have more comprehensive suggestions due to all being part of the same company, and due to general trends of anthropomorphism in robots [13]. An initial survey containing 41 questions, representing 6 robot scenarios, 3 comparisons, one demographics question, and one general question about robots in the workplace was sent to the full set of company employees, comprising 250+ employees. Participation was voluntary and anonymous. This section describes the contents of this survey and the grounded coding methods used to analyze the data.

	Interaction Partner	**Obstacle**
Human	Scenario 1,	Scenario 2, Fig. 3b
Robot	Scenario 3, Fig. 3c	Scenario 4, Fig. 3d
Box	Scenario 5, Fig. 3e	Scenario 6, Fig. 3f

Table 2. Six scenarios that participants were asked about in the company survey.

The survey used a scene-based approach to solicit employee design ideas for improved robot expressions. In other words, the participants were presented with a situation and asked what they would like to happen if they were in the presented situation. The aim of using the scene-based approach was to get employee participants to make specific suggestions based on the scenario, instead of telling us vaguely how the robot could communicate better overall. As past work has shown, scenario-based thinking can help participants come up with concrete suggestions for interaction modes or actions. To develop these scenarios, we used a combination of several methods to ensure that they were representative of scenarios in which the robots do not currently communicate, but employees would like them to. First, we used participant observation by watching the robots on the test floor to determine what situations the robots were commonly in. In addition to participant observation, we also conducted informal interviews with employees on the testing team to see when they got frustrated with the robot's lack of communication. Thus, six navigation scenarios were populated that were identified as common to factory robot operations (Fig. 3).

To collect employee input on communication needs and channels, each employee saw all six scenarios one at a time in a randomized order. Each scenario was followed by a group of questions. In order to prevent leading the responses, each of these question groups began with an open-ended question, then asked for specific recommendations about robot movement, lights, or sound, and suggestions for their implementation:

1. **What should the robot do next?**
2. **How could the robot move next in this context?**
3. **How could the robot use lights in this context?**
4. **How could the robot use sounds in this context?**

In total, 65 employees responded to the survey. Departments represented included Autonomy, Product Applications, Client Success, Field Services, Operating Systems, Test Team, Sales, Electrical Design, Mechanical Design, User Experience, and Management. This represents a broad range of both technical and non-technical backgrounds in the company.

(a) **Scenario 1**: “You are standing in a large open space and **a robot approaches you to deliver supplies.**”

(b) **Scenario 2**: “You are standing in a large open space and **a robot approaches you and needs to get by.**”

(c) **Scenario 3**: “You are standing in a large open space and observe two robots. **The white robot is delivering supplies to the purple robot.** The purple robot is parked, but able to move if desired.”

(d) **Scenario 4**: “You are standing in a large open space and observe two robots. **The white robot needs to get by the purple robot** The purple robot is parked, but able to move if desired.”

(e) **Scenario 5**: “You are standing in a large open space and observe a robot. **The robot is delivering supplies to the box.**”

(f) **Scenario 6**: “You are standing in a large open space and observe a robot. **The robot needs to get by, but there is a box in the space.**”

Fig. 3. Survey Scenarios, with the robot direction is indicated with an arrow and the robot goal is indicated by red, bold font. The white robot is the subject of the question. The purple robot is an outside robot in the scene.

To analyze the resulting survey data, the research team used grounded coding [9] to develop annotations, categorize common themes, and identify significant concepts for robot communication development and assessment. The first step of grounded coding is for a researcher to develop "codes" a.k.a. annotations for information that comes up related to their research question. As new annotations are developed, a researcher can return to previously labeled data as an iterative process. These annotations are then grouped into common categories or key words, representing meta-concepts that were present in the data. These meta-concepts can be defined and evaluated with independent reviewers to confirm the theory building resulted in repeatable concepts.

In this work, we conducted this annotation, theory building, and evaluation process for three channels of factory robot communication: light, sound, and motion. Our evaluation of the data presented three themes: (1) What are situations in which robot communication is desired? (2) What communication strategies are most appropriate for these situations? (3) What are the styles of communications suggested for each channel?

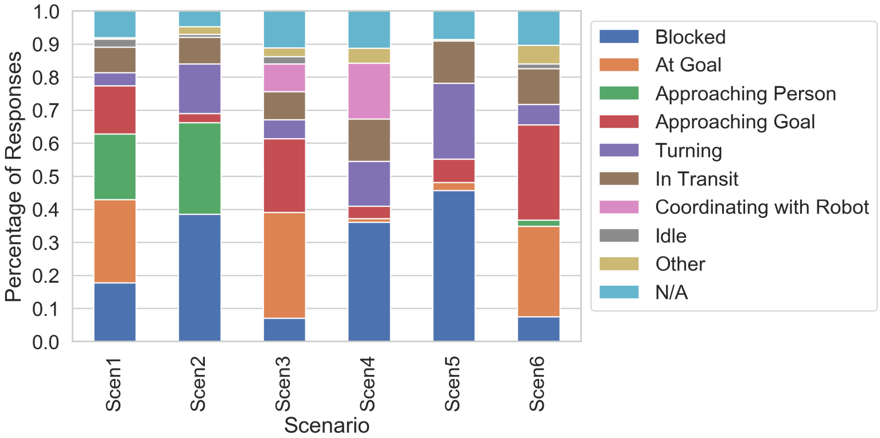


Fig. 4. Percentage of responses labelled with each Situation code by scenario.

All annotations and coding were done at once for each response. First, the questions were grouped by scenario. The first question in the scenario coding was done for communication situations and desired channels. All responses in the first question were coded before moving on to the second question. Question two was coded for communication situation, motion channel, and motion signals. Again, all responses in question two were coded before moving on to question three. All responses to the third question were coded for communication situation, light channel, and light signals. Finally, all responses to the fourth question were coded for communication situation, sound channel, and sound signals. Once all four questions were coded in one scenario, the coder moved onto the next scenario and repeated the process. Intercoder reliability was done on half the data, specifically scenarios 1, 4, and 5.

Overall, the company survey intercoder reliability was 91.8%

4.1 Category 1: Situations Important for Communication

4.1.1 Analysis Methods. First, the coder annotated for verbs describing robot actions, the objects of those verbs if they existed, and adjectives describing the robot's state. Then, the annotations and the scenario were used to determine under which codes the response fell. Responses could fall under multiple codes and be marked as such. Responses could also fall under a N/A code if the response did not mention any communication situations. The codes and their description are seen in Table 3.

4.1.2 Results. The percentage of responses in each code for each survey scenario is presented in Figure 4. In Scenario 1, the most frequent code labelled was the "At Goal" code. The second most popular was "Approaching Person," and the third most popular was "Blocked," followed by "At Goal." For Scenario 2, the most frequent code labelled was the "Blocked" code. Like Scenario 1, the second most popular was "Approaching Person." Third most popular was "Turning." Scenario 3 had the same top two as Scenario 1The top response for Scenario 4 was the "Blocked" code. Next most frequent was "Coordinating with Robot," followed by "Turning" and "In Transit." Like Scenario 4, for Scenario 5 the most frequent code labelled was the "Blocked" code. The second frequent was "Turning," which was followed by "In Transit." In Scenario 6, the most frequent code labelled was the "Approaching Goal" followed by "At Goal," and thirdly followed by "In Transit."

Category	Code	Criteria
Situation	*Blocked*	- Uses the word blocked to describe the robot's state - Mentions that something is blocking the robot - Mentions the robot should wait for something to move out of the way - Mentions the robot should indicate that something is in the way - Mentions the robot going around an obstacle Or similar language
	At Goal	- Mentions the robot is in a safe state or "WIP" mode - Mentions the robot is arriving at the goal or target - Mentions the robot is ready for delivery - Mentions the robot is stopped at a safe distance from the target - Mentions the robot being at the final position or destination Or similar language
	Approaching Person	- Mentions the robot taking an action or signalling because it is near a person - Mentions indicating to a person if that person is a bystander (ie, the person is not the goal) - Mentions "you" as the object of the verb approach - Mentions moving towards a person Or similar language
	Approaching Goal	- Mentions the robot taking an action or signalling because it is near, but not yet at, the goal - Mentions docking with the robot as the subject of the verb dock - Mentions the final approach Or similar language
	Turning	- Mentions turning with the robot as the subject of the verb turn - Mentions rotating with the robot as the subject of the verb rotate Or similar language
	In Transit	- Mentions driving without any additional specification - Mentions forwards motion without any additional specifications - Mentions navigating without any additional specifications Or similar language
	Coordinating with another Robot	- Mentions coordinating in relation to another robot - Mentions working with another robot - Mentions synchronizing with another robot Or similar language
	Idle	- Mentions the robot does not have a task - Mentions the robot is idle - Mentions the robot is waiting for a task or job Or similar language
	Other	Situation that does not fit in any of the above categories
	N/A	Does not mention situation

Table 3. Situation codes for the Company Survey grounded coding

Category	**Code**	**Criteria**
Channel	*Yes [Motion / Lights / Sounds]*	- The response mentioned they wanted the robot to communicate using [motion / lights / sounds].
	No [Motion / Lights / Sounds]	- The response mentioned they DID NOT want the robot to communicate using [motion / lights / sounds].
	N/A [Motion / Lights / Sounds]	- The response did not mention communicating with [motion / lights / sounds].

Table 4. Desired channel codes for the Company Survey grounded coding.

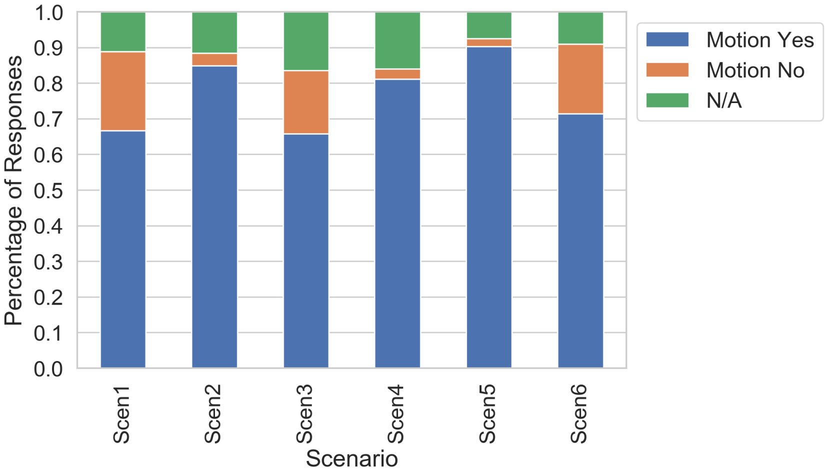


Fig. 5. Percentage of responses labelled with each Motion Signal code by scenario.

The intercoder reliability for scenario 1 was 89.9%, scenario 4 was 91.9%, scenario 5 was 91.7%, the overall Situation Category intercoder reliability was 91.2%.

4.2 Category 2: Desired Communication Channels

4.2.1 Analysis Methods. The coders annotated if the response mentioned using either motion, lights, or sound to communicate. The coder also annotated if the response said specifically to *not* use any or all of the channels. Using the annotations, the response was categorized as "[Motion/Lights/Sound] Desired," "NO [Motion/Lights/Sound] Desired," or "N/A [Motion/Lights/Sound]." The N/A category was again used when the response did not mention either using or not using a specific channel. These codes can be seen in Table 4.

4.2.2 Results. The intercoder reliability across all desired channels was 94.3%. For the motion channel, intercoder reliability for scenario 1 was 91.0%, scenario 4 was 90.7%, scenario 5 was 91.0%, and the overall motion channel intercoder reliability was 90.9%. Across all scenarios, over 60% of responses wanted the robot to use motion, with Scenario 2, 4, and 5 being the most popular for the motion channel, as seen in Figure 5.

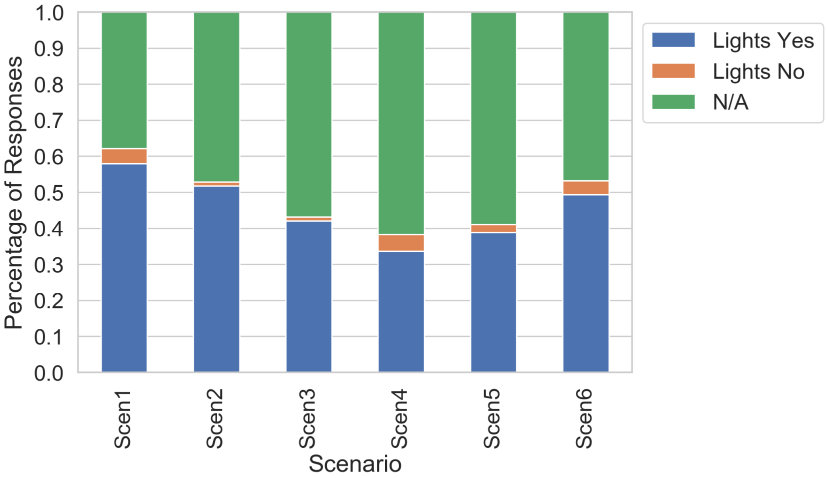


Fig. 6. Percentage of responses labelled with each Light Channel code by scenario.

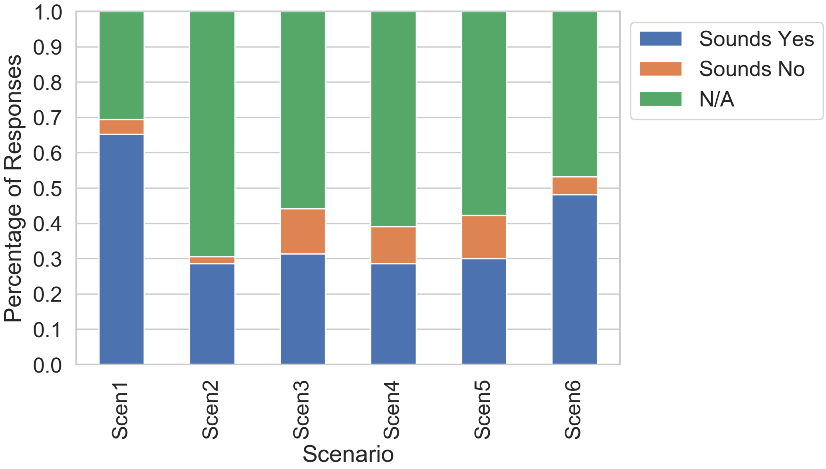


Fig. 7. Percentage of responses labelled with each Sound Channel code by scenario.

For the lights channel, intercoder reliability for scenario 1 was 95.8%, scenario 4 was 94.4%, scenario 5 was 91.9%, and overall intercoder reliability was 94.0%. Lights were most desired in Scenarios 1 and 2 where there were people involved and in Scenario 6 in which the robot was blocked by an obstacle, as seen in Figure 6.

The intercoder reliability for the sounds channel was 92.8% for scenario 1, 99.1% for scenario 4, 98.0% for scenario 5, and 97.1% overall. Sounds were most desired in Scenarios 1 and 6, where the robot is delivering something to a human and where the robot is blocked by an obstacle respectively. Scenarios 2 through 5 were less popular for the sound channel, with less than 40% of responses wanting sound in, as seen in Figure 7.

Category	**Code**	**Criteria**
Motion Signal	*Functional*	- The motion desired is purely functional and not meant to be communicatory - Only mentions necessary movements - Uses words like "direct," "minimal," and "efficient" to describe the motion
	Expressive	- The motion desired is intended to communicate or convey something beyond being functional - Mentions emotion adjectives when describing desired movement - The motion desired is not necessary to complete the task
	N/A	- The response did not mention desired motion signals

Table 5. Motion signal codes for the Company Survey grounded coding.

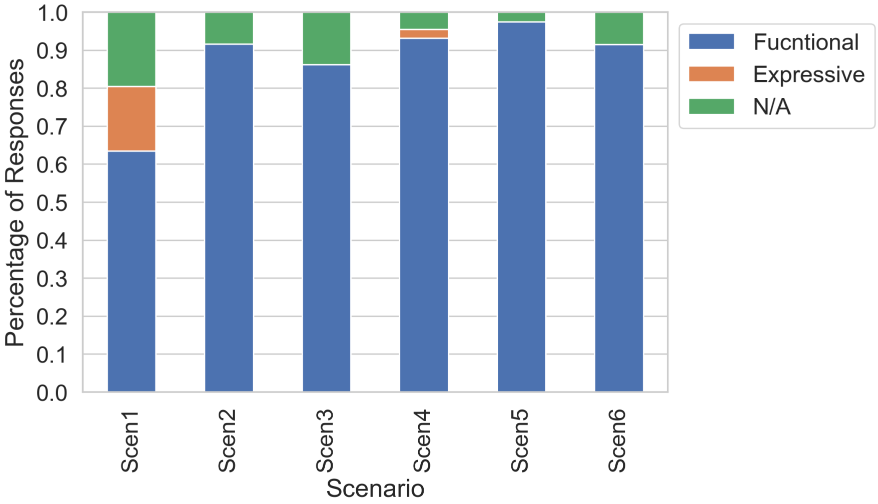


Fig. 8. Percentage of responses labelled with each Motion Signal code by scenario.

4.3 Category 3: Desired Communication Signals

4.3.1 Analysis Methods: ***Motion.*** The coders first annotated all adjectives describing the motion speed, adjectives describing the general motion, and phrases describing the motion. The all verbs used to describe the movements were annotated. Next, any adjectives or adverbs describing the motion direction or path, like “straight” or “curved” were annotated. The coders then annotated for any specific movement phrases. The annotations lead to the codes seen in Table 5.

4.3.2 Results: ***Motion.*** For motion signals, the intercoder reliability for scenario 1 was 92.9%, scenario 4 was 95.6%, scenario 5 was 92.9%, and the intercoder reliability overall for motion signals was 93.8%. Across all scenarios, the “Fucntional” code always had the highest percentage of

Category	Code	Criteria
Light Signal	*Sweeping / Directional*	- Mentions the lights "pointing" to or "indicating" something of interest - Mentions the lights "drawing" the robot or "pulling" towards something - Mentions a "sweeping" motion of the lights Or similar language
	Vehicle	- Mentions taking "inspiration" or "cues" from traffic, vehicles, or the road - Uses a metaphor or simile comparing the robot to a car - Uses language typically used when referring to vehicles, such as "high beams" or "headlights" Or similar language
	Eyes	- Uses a metaphor or simile comparing the robot lights to eyes - Mentions the robot should "look" at something Or similar language
	Other	- Does not fit into any of the above codes or is too vague to be included in the above codes
	N/A	- Does not mention light signals

Table 6. Light signal codes for the Company Survey grounded coding.

responses, however it was significantly less popular in Scenario 1 with 63.4%, than the other five scenarios.

4.3.3 Analysis Methods: ***Lights****.* The coders annotated all adjectives describing the colors of the lights and if the color is specified for a certain state. All other non-color adjectives describing the lights were then annotated. Verbs describing the movement of the lights were annotated, and if that verb had an object, the object was noted. Lastly, the coders annotated any metaphors and similes used to describe the lights. The codes that arose from these annotations can be seen in Table 6.

4.3.4 Results: ***Lights****.* For the lights signals, intercoder reliability for scenario 1 was 90.9%, scenario 4 was 86.5%, scenario 5 was 93.7% and the overall intercoder reliability was 90.2%. Across all scenarios, the "Other" code always had the highest percentage of .Vehicle lights were most popular in Scenarios 1 and 2, which both involved people. Sweeping lights were most desired in Scenarios 1 and 3, when the robot is delivering to a human and to a robot respectively. Eyes were only mentioned in Scenarios 2, 4, and 6. These results can be seen in Figure 9.

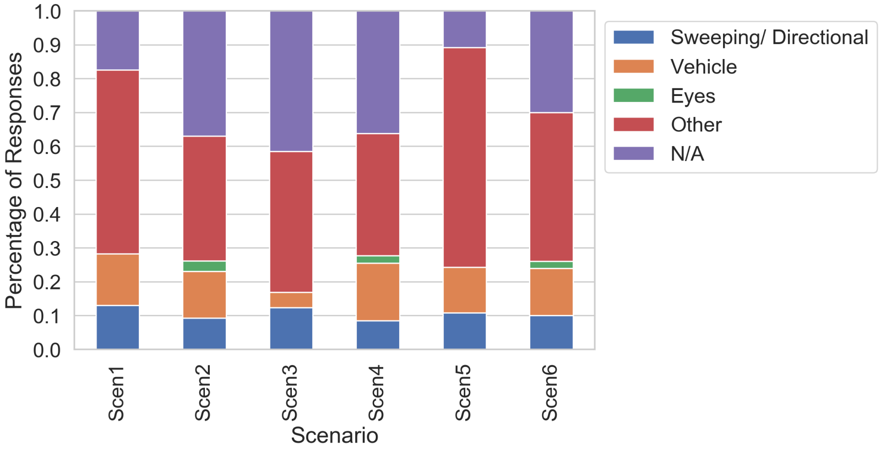


Fig. 9. Percentage of responses labelled with each Light Signal code by scenario.

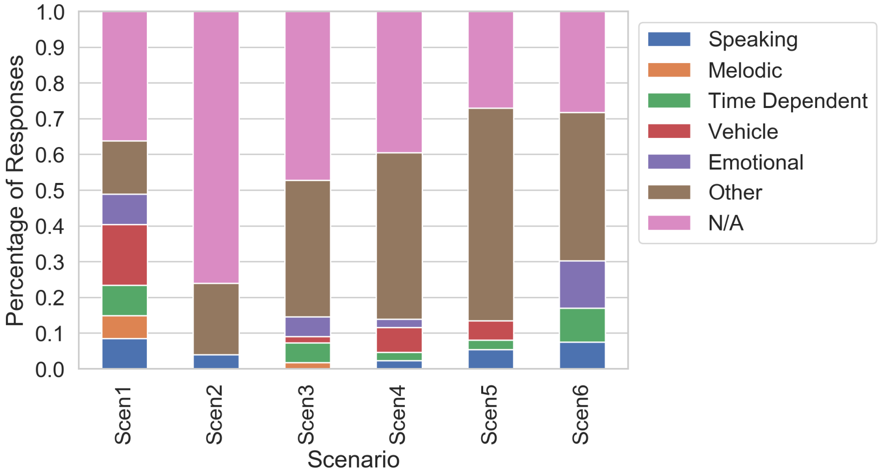


Fig. 10. Percentage of responses labelled with each Sound Signal code by scenario.

4.3.5 Analysis Methods: ***Sound****.* First the coders annotated all adjectives describing the tones or pitch of the sounds, emotional adjectives describing the sounds, and adjectives and phrases describing the volume. Then all phrases describing sound frequency were annotated. Next, the coders annotated any verbs describing the sounds and if the verbs have a specified situation, such as honking when blocked, the situation was annotated too. Then any metaphors and similes used to describe the sounds were annotated, and finally the coders annotated any specific melodies, musical phrases, or sonic phrases mentioned. From these annotations, the responses were coded following the codes listed in Table 7.

4.3.6 Results: ***Sound****.* For the sound signals, intercoder reliability for scenario 1 was 86.9%, scenario 4 was 97.0%, scenario 5 was 98.3%, and overall sound signal intercoder reliability was 93.8%. In Scenario 1, which is the robot delivering something to a human, the most desired sound signal

Category	Code	Criteria
Sound Signal	*Melodic*	- Mentions a specific melody or song Or similar language
	Time Dependent	- Mentions volume or frequency increasing or decreasing over time - Uses the phrase "over time" referring to the sound Or similar language
	Speaking	- Mentions the robot should "say" something or "talk" - Specifies the robot should use words Or similar language
	Vehicle	- Uses a metaphor or simile to compare the robot to a vehicle in regards to sound - Mentions vehicle-specific noises like "honking the horn" or "turn signal clicking" - Mentions taking "inspiration" or "cues" from traffic or the road Or similar language
	Emotional	- Uses emotion adjectives when referring to sound - Says the robot should use sound to "show it is [emotion]" Or similar language
	Other	- Does not fit into any of the above codes or is too vague to be included in the above codes
	N/A	- Does not mention sound signals

Table 7. Sound signal codes for the Company Survey grounded coding.

was a vehicle sound. Apart from Scenario 1, the most popular sound signal was the other category. Speaking sounds, time dependent sounds, and emotional sounds were most seen in Scenarios 1 and 6 where the robot is delivering to a human or blocked by an obstacle respectively, which both require human action. Melodic sounds were only mentioned in Scenario 1 and 3, which are the robot delivering to a human and to an object, respectively. These results can be seen in Figure 10.

4.4 Overall Results

Overall the most popular scenarios were “Blocked,” “At Goal,” “Approaching Goal,” “Turning,” and “In Transit.” In desired channels, the most popular was motion, followed by lights then sound. However, if only the responses that said Yes or No are considered (not “N/A”) then the lights channel had the highest desired percentage with 93.9% wanting lights, followed by motion with 87.8% desiring motion, followed by sounds with 82.5% desiring sound. For motion signals, by far the most popular was purely functional motion. For lights, many answers were too vague to fit into a particular category, but of those that did vehicle lights were the most popular followed by sweeping. Similarly for sound signals, many responses were too vague to fit a particular category, but of those that did the most popular were emotional sounds, followed by time dependent and vehicle sounds, followed by speaking sounds.

5 TECHNOLOGY IMPLEMENTATION: EXPRESSIVE FACTORY LIGHTS

Using the results of the company survey and prior work in social robotics, custom light sets were created and implemented on the robots in such way that they seamlessly integrated into the full software stack.

5.1 Different Communication States

Based on the company survey results, participant observation, and sensing limitations, four codes from Cateogry 1: Situations Important for Communication were chosen to become the “states” for creating custom lights: “blocked,” “at goal,” “turning,” and “idle,” as seen in Figure **??**. These states were determined combining sensor information about the location of the robot, the motion of the robot, the goal location (if one exists), and any points blocking the robot (if they exist). The “blocked” state was triggered when the robot was not at the goal, had not moved within a certain distance within a certain amount of timesteps, and had detected points blocking the robot on the lidar. The “at goal” state was triggered when the robot was not moving and within a certain distance of the goal. The “turning” state was triggered when the robot path had a radius of curvature within a certain range. This was to avoid triggering the “turning” state when the robot was just on a curved path. The “idle” state was triggered when the robot did not have any goals in its queue.

5.2 Technical Implementation and Integration

To ensure that the custom light sets could be easily integrated into the current lights, they were created using a similar process, with some additional automation to make them easier to change and be dynamic. The standard lights for the OTTO 100 are set using PNG files, with the horizontal axis being lights around the light strip and the vertical axis being time. Dymanic PNGs that could be read the same way as the original light PNGs were created. Using this method, the new lights were able to be easily added to the software stack without interrupting any other processes on the robot. The software flow and integration can be seen in Figure 11.

5.3 Why these lights?

Three custom LED sets were created for the user study: Car-like, Sweeping, and Heartbeat. The Standard lights were used as a control, to which the three custom light patterns were compared. All four light sets can be seen in Figure 12. The Car-like and Sweeping light patterns were based on the results of the company survey. The Heartbeat pattern was inspired by minimal light communication with colors, as seen in the work of Kim Baraka, which showed that red pulsing light is the best combination for indicating a blocked state [3]. Work by Song, et al, also showed that red lights were often viewed as hostile when shown in a pulsing pattern, which may translate well to the “blocked” state, as “blocked” is a negative state for the robot to be in [41]. The colors for the states in the custom set were based on the company survey. Many participants suggested red for “blocked” and green lights for “at goal.” Turning was kept white because it is a neutral color.

The Standard lights that already existed on the robot for the desired states only consisted of two light patterns: “turn” lights, and static lights for “drive,” “blocked,” “at goal,” and “idle.” The static lights consisted of white lights on the two front corners, and two red lights on back two corners, as seen in Figure 13a. The lights for the “turn” state can be seen in Figure 13b.

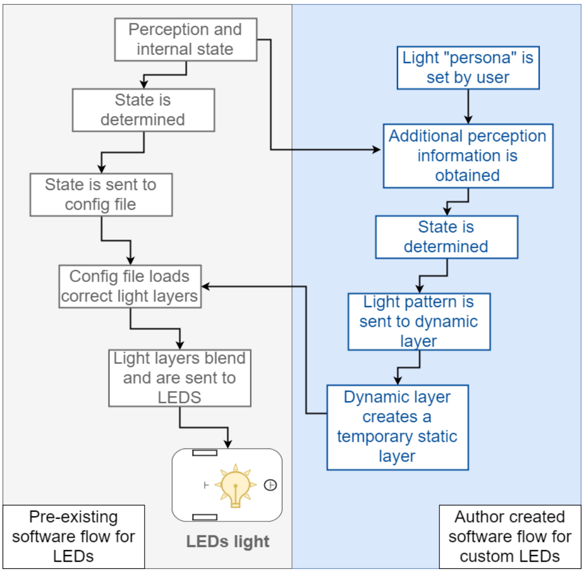


Fig. 11. We extended existing software for setting the 340 multicolor-LED strips encircling the robots (on left) to allow for the creation of dynamic, custom LED sets (on right).

The Car-like lights were inspired by the company survey results and were based on the analogy of a car, with off-white headlights in the front and dimmer red lights in the back in the “drive” state. For “blocked,” “turning,” and “at goal” states, the rear and headlights flashed in various colors. The flashing was chosen to follow the car analogy, as the main way cars communicate with lights is through the flashing or blinking of the headlights and rear lights. When blocked, the headlights flashed bright white, to match the car analogy, and the back headlights flashed red, to match the red color present in the “blocked” state in the other custom sets. When in the “turning” state, the headlight and rear light on the direction of turning flashed white, to keep to the car analogy and match the white color for turning of the other custom sets. When in the “at goal” state, the headlights and rear lights flashed green, to match the colors of the other custom sets. An example of the Car-like lights can be seen in the “turn” state in Figure 14a.

The Sweeping lights were based on the results of the company survey. The intention behind them was to have the lights move around the robot and converge to a particular point of interest, whether that be the goal or an obstacle. The pattern was created with dashed lines so that it would be easier to see the movement of the lights; with a solid shrinking light, it can look static from certain angles if you cannot see the ends which are shrinking. However, with moving dashed lines you can see the movement and direction from any angle. A similar light pattern can also be seen in the work of Song, et al, which was shown to be attractive and likable to humans interacting with the robot [41]. The “blocked” lights were red and converged to the obstacle, the “turning” lights were white, and the “at goal” lights were green. An example of the Sweeping lights can be seen in the “at goal” state in Figure 14b.

The Heartbeat lights were based on the work of Kim Baraka [3–5], having the lights pulse from dim to bright on all the lights around the light strip. Additional works utilizing pulsing lights include work by Song, et al. on a static (non-moving) robot, and by Harrison, et al. on a single LED [16, 41]. Having the light pattern around the whole robot is simple, with minimal movement, and visible from all sides of the robot. The “blocked” lights were red, the “turning” lights were white,

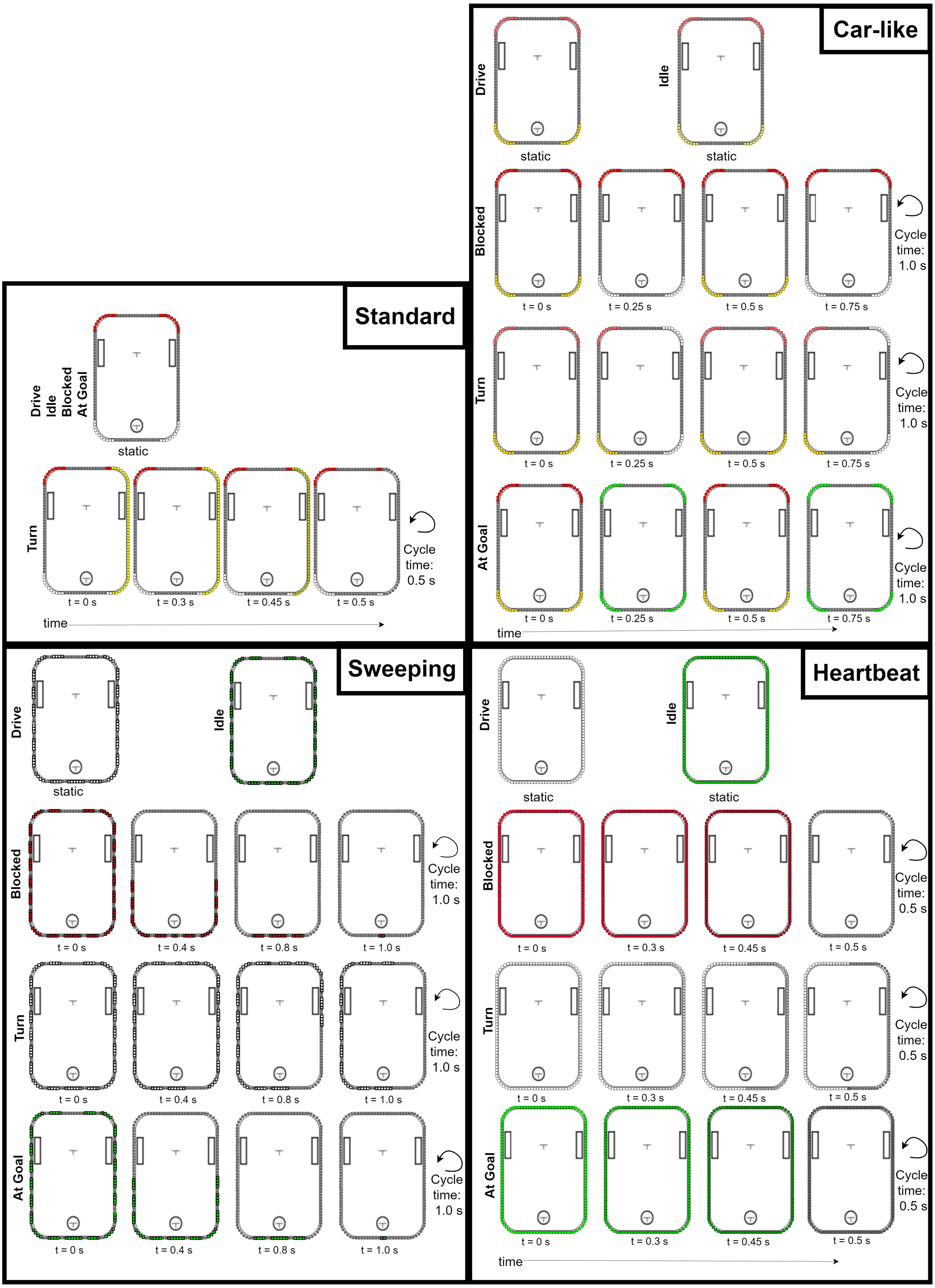
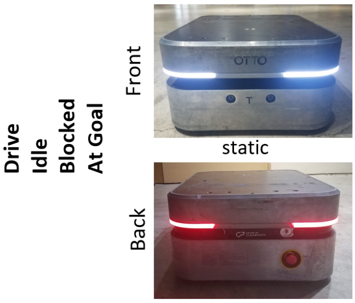


Fig. 12. Pre-existing lights (Standard) and newly designed custom light sets (Car-like, Sweeping, Heartbeat).



(a) Standard lights in the “drive,” “blocked,” “at goal,” and “idle” states on the OTTO 100.

(b) Standard lights in the “turn” state on the OTTO 100.

Fig. 13. Standard light set for the OTTO 100

and the “at goal” lights were green. An example of the Heartbeat lights can be seen in the “blocked” state in Figure 14c.

6 USER STUDY: EXPRESSIVE LIGHT EVALUATION AND RESULTS

A user study was created to evaluate the standard and custom light sets for legibility. First is an overview of the study, followed by the study procedure, participant demographics, hypotheses and finally, results.

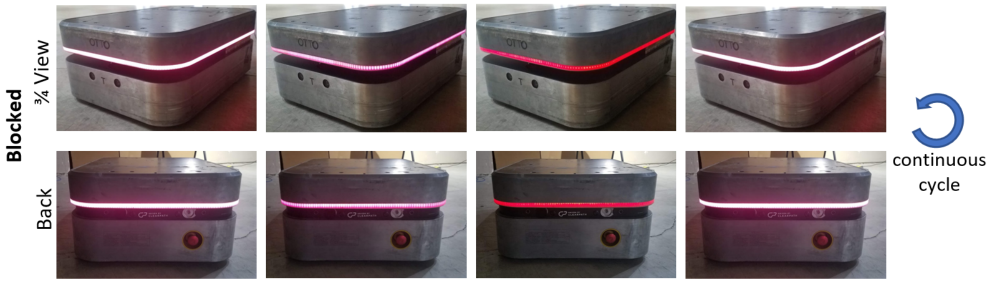
6.1 User Study Overview

The goal of this study was to evaluate the relative merits of the four light set patterns, described in detail in Section 5.2, across four different robot communication states, described in detail in Section 5.1. These variables and all their combinations can be seen in Table 8. For this purpose, participants were invited to the Clearpath test floor, and took part in a human-robot path sequence mimicking a problematic customer site described below. The experiment was within-participants, in other words, all participants experienced all four light sets. Each participant saw the Standard light set first, then saw the three custom light sets in a random order. After each example they responded to survey questions exploring the legibility of the light sets by state, and their subjective experiences of the lights with regards to safety, comfort and distraction.

The backdrop of the study took inspiration from a common problem at a specific customer site involving an L-junction and several moving entities backing in and out of the space (Fig. 15). In the customer site, forklifts would occasionally block the OTTO 100s, and because of their respective dropoff locations, the OTTO 100s would occasionally block the forklift drivers also. This scenario was chosen because understanding robot state well in an area with occlusions and the potential for

(a) Car-like lights in the “turning” state on the OTTO 100.

(b) Sweeping lights in the “at goal” state on the OTTO 100.



(c) Hearbeat lights in the “blocked” state on the OTTO 100.

Fig. 14. Examples of the custom light sets on the OTTO 100

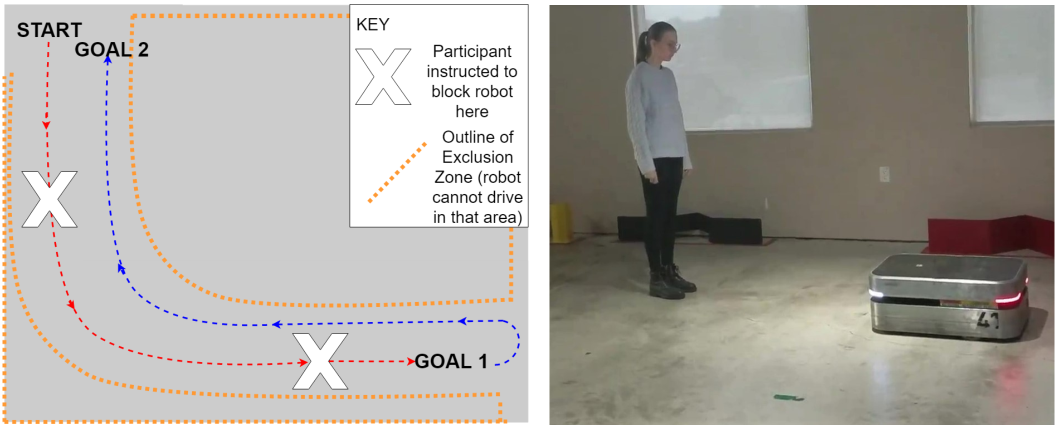
multiple actors would be critical for rapid and successful coordination. To mimic this scenario, a path was set up for the robot to drive in a large L-shape, and the user study participant blocked the robot twice along that path. To ensure that the robot did not stray from this L-shaped area, two Exclusion Zones were created so that the robot could not enter the large free space surrounding the desired paths. To drive from START to GOAL 1 and GOAL 1 to GOAL 2, the study facilitator created waypoints at each of these places, told the robot to move to them at the appropriate time, and the OTTO 100 moved autonomously between them.

6.2 Study Procedure

First, the participant came to the question table, as seen in Figure 15, and filled out a short demographic survey. Then, they were told they would see four different light patterns and to only focus on the lights in the study, not the motion. The sounds were muted on the robot for this experiment.

		Light Set			
		Standard	*Sweeping*	*Car-like*	*Heartbeat*
State	*Blocked*	N/A	Red dashed lights moving on both sides of the robot, converging at blocking obstacle	Front yellow lights blink white	Red light around the entire robot pulsing
	At Goal	N/A	Green dashed lights moving on both sides of the robot, converging at the center front	Front yellow lights and back red lights blink green	Green light around the entire robot pulsing
	Turning	Yellow light pulsing on direction of turn	White dashed lights moving from center front to center back on side of turn	Front yellow light and back red light blink white on side of turn	White light on side of turn pulsing
	Idle	N/A	Green dashed lights around the entire robot static	Front yellow lights and back red lights dimmed	Green light around the entire robot static

Table 8. Variables and descriptions of each variable combination for the in-person user study to evaluate the custom light sets and custom states.



(a) User study layout diagram (b) Staged photo of a participant blocking the robot

Fig. 15. User study layout: The robot moved from the START to GOAL 1 on the red path, during which it was blocked by the participant twice, as seen in subfigure 15b. The robot then moved from GOAL 1 to GOAL 2 along the blue path while the participant observed. The dotted orange line marks the boundaries of the Exclusion Zones, which were set to ensure that the robot stayed within the L-shaped user study area.

They were then told that the robot would follow the L-shaped and where and when to block the robot, as seen in in Figure 15.

The participants were shown the standard OTTO 100 lights first as a baseline, which was done so that participants would compare the custom light sets to the pre-existing Standard lights. After seeing the first set of lights, the participant returned to the question table, where they filled out some online questions and had a short interview with the person running the study. The three sets of custom lights were seen in a different order, randomized for each participant. After each set of lights, the participant came back to the question table to fill out online questions and have a short interview. After all the light sets were viewed, a short closing interview was conducted.

6.3 Participant Demographics

At the time of this work, Clearpath robots were deployed at over 100 factories, all with diverse setups and end user interactions [10]. As previously mentioned in Sections 3 and 3, we hypothesized that using employees from diverse subteams at Clearpath as participants would provide: (1) a nuanced view of how people with different expertise view robot light communication and/or (2) show that despite different background knowledge of the robots, that most people have similar ideas about robot light communication. Both of these outcomes would allow for a more comprehensive evaluation of the custom lights.

Participants were recruited via email to the cross-company list. Participation was completely voluntary, and participants signed up on a calendar to participate. As part of our demographic survey, we asked participants about which team they worked for and their interactions and familiarity with the robots. Overall, there were 30 participants in the study. As part of our demographic survey, we asked about their interactions and familiarity with the robots. 23.3% of the participants interacted with the robots daily, 43.3% interacted with the robots a few times a week, 13.3% interacted with the robots weekly, 13.3% interacted with the robots monthly, and 3.3% of participants interact with the robots less than once a month. 16.7% of participants knew all the Standard light signals and when they occurred, 30.0% of participants knew most of the Standard light signals and when they occurred, 33.3% of the participants knew some of the Standard light signals and when they occurred, 16.7% of participants knew some of the Standard light signals but not what they meant, and 3.3% of participants did not know any of the Standard light signals or what they meant. Additionally, we asked which team the participants worked for, with 30% working in software, 10% in hardware, 16.7% in testing, 23.3% in product design, 10% in client services, and 10% in marketing.

6.4 Hypotheses

For signal clarity, i.e., evaluating which light pattern is best for which state, it was hypothesized that Sweeping lights would be best for “blocked” because the lights point towards the obstacle, showing where the robot is blocked, while grabbing the attention of the user with the moving lights and color gradients. Car-like lights were hypothesized to be best for “turning” because the lights were similar to a turn signal on a car and were easily recognizable. Heartbeat lights were hypothesized to be best for “at goal” because they were simple and were less likely to draw unwanted attention, but show the task is complete. It was also hypothesized that Car-like lights may be best for “idle” since the “idle” lights like “drive” lights but dimmer to indicate not active.

It was hypothesized that the employee teams will significantly change the ratings only when there is a difference in how frequently participants on that team interact with the robots and are familiar with the lights. To this end, it was hypothesized that those who interact with the robots frequently and those are familiar with the light patterns will like the Car-like lights best and think they have the most signal clarity because this light set is closest to the current Standard light set these participants are used to seeing.

6.5 Results

Each set of data for Signal Clarity (6.5.1) and Emotional Response (6.5.2) was tested to see if it was a normal distribution using a histogram. All data sets were non-normal distributions so to determine significance, the non-parametric Kruskal-Wallis test was done. If there was significance, the data was then tested pairwise using both the Mann-Whitney test and Tukey HSD test as post-hoc tests to determine pairwise significance.

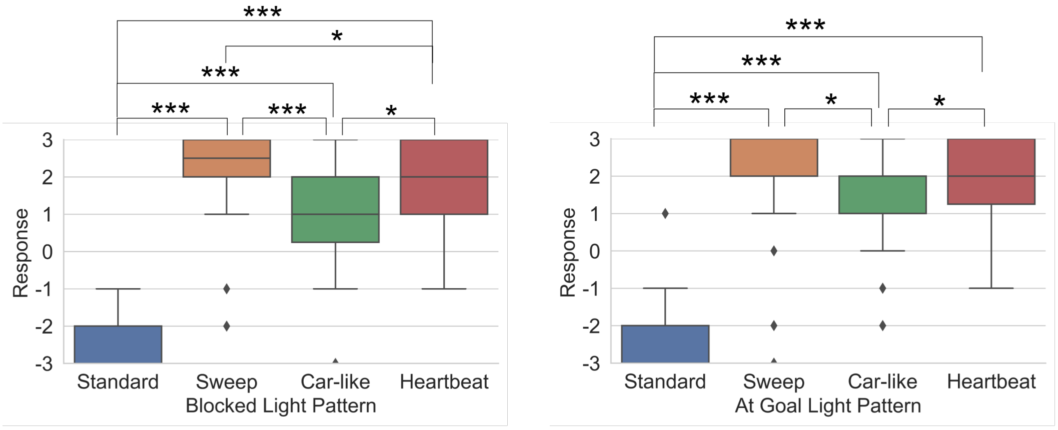
6.5.1 Signal Legibility. For all responses on signal legibility, the responses ranged from -3 (the signal was very unclear) to 3 (the signal was very clear).

For the "blocked" light set, the Sweeping was rated the clearest signal (median = 2.5, σ =1.17), followed by the Heartbeat set (median = 2.0, σ = 1.83), and then the Car-like set (median = 1.0, σ = 1.53). The Standard was rated the least clear signal (median = -3.0, σ = 0.56). LED light set (including the Standard) also very significantly impacted clarity in signalling "blocked" state (F(3,115) = 73.84, $p < .001$) using Kruskal-Wallis test. The Mann-Whitney U test was used to compare the custom light sets to the Standard light set, as seen in Figure 16a. The Sweeping set was rated significantly more legible than the Standard set (U = 6.0, $p < 0.001$), as did the Car-like set (U = 22.5, $p < 0.001$) and the Heartbeat set (U = 1.0, $p < 0.001$). The Mann-Whitney U test was also used to test significance between the custom sets showing the Sweeping set was rated significantly higher than the Car-like set (U = 225.5, $p < 0.001$). The tests also showed weaker significance between the Sweeping set and the Heartbeat set (U = 344.5, $p = 0.49$) and the Heartbeat set and the Car-like set (U = 317.0, $p = 0.21$). The Tukey HSD test confirmed all significant pairs.

For the "at goal" light set, the Sweeping set (median = 2.0, σ =1.40), the Car-like set (median = 2.0, σ = 1.56), and the Heartbeat set (median = 2.0, σ = 1.01) performed similarly. The Standard was rated the least clear signal (median = -3.0, σ = 0.98). LED light set (including the Standard) also very significantly impacted clarity in signalling "at goal" state (F(3,115) = 65.14, $p < 0.001$) using Kruskal-Wallis test. The Mann-Whitney U test was used to compare the custom light sets to the Standard light set, as seen in Figure 16b. The Sweeping set was rated significantly more legible than the Standard set (U = 34.0, $p < 0.001$), as did the Car-like set (U = 42.0, $p < 0.001$) and the Heartbeat set (U = 7.0, $p < 0.001$). Significance between the custom sets was also tested using the Mann-Whitney U test. Weaker significance was found between the Sweeping set and the Car-like set (U = 330.5, $p = 0.032$) and the Heartbeat set and the Car-like set (U = 307.0, $p = 0.014$). The Tukey HSD test confirmed all significant pairs.

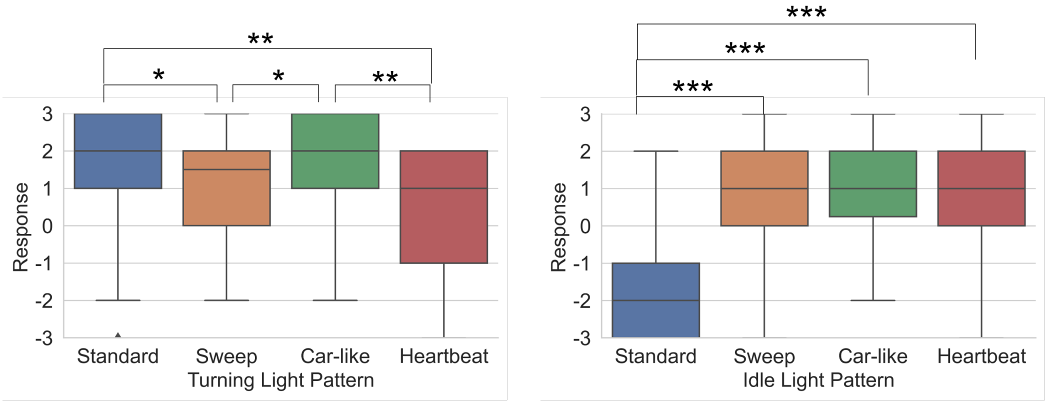
For the "turning" light set, the Car-like set (median = 2.0, σ =1.24) and the Standard set (median = 2.0, σ = 1.80) performed the best, followed but the Sweeping set (median = 1.5, σ = 1.54). The Heartbeat set was rated the least clear signal (median = 1.0, σ = 1.60). LED light set (including the Standard) significantly impacted clarity in signalling "turning state" state (F(3,115) = 11.82, $p = 0.008$) using Kruskal-Wallis test. The Mann-Whitney U test was used to compare the custom light sets to the Standard light set, as seen in Figure 16c. The Standard set was rated significantly more legible than the Heartbeat set (U = 284.0, $p = 0.006$), and significantly more legible than the Sweeping set (U = 326.0, $p = 0.03$). Significance between the custom sets was also tested using the Mann-Whitney U test. The Car-like set was significantly better than bot the Sweeping set (U = 306.5, $p = 0.014$) and the Heartbeat set (U = 264.5, $p = 0.002$). The Tukey HSD test confirmed all significant pairs.

For the "idle" light set, the Sweeping set (median = 1.0, σ =1.53), the Car-like set (median = 1.0, σ = 1.34), and the Heartbeat set (median = 1.0, σ = 1.74) performed similarly. The Standard was rated the least clear signal (median = -2.0, σ = 1.68). LED light set (including the Standard) significantly impacted clarity in signalling "idle state" state (F(3,115) = 36.04, $p < 0.001$) using Kruskal-Wallis test. The Mann-Whitney U test was used to compare the custom light sets to the Standard light set, as seen in Figure 16d. The Sweeping set was rated significantly more legible than the Standard set (U = 129.0, $p < 0.001$), as were the Car-like set (U = 103.5, $p < 0.001$) and the Heartbeat set (U = 154.0 p



(a) “These lights communicated clearly when the robot was **blocked**.”

(b) “These lights communicated clearly when the robot was **at goal**.”



(c) “These lights communicated clearly when the robot was **turning**.”

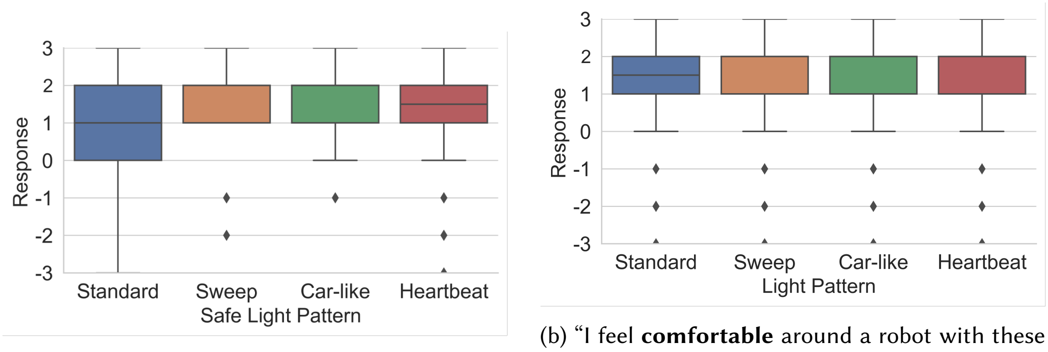
(d) “These lights communicated clearly when the robot was **idle**.”

Fig. 16. Survey responses to "These lights communicated clearly when the robot was **blocked, at goal, turning, idle**". Responses are on a seven point Likert scale from -3 (I strongly disagree with this statement) to 3 (I strongly agree with this statement). Shown as median, 25% quarantile, 75% quarantiles, and outliers by light pattern. Significance is shown as * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$.

< 0.001). There was no significance between the custom sets as tested using the Mann-Whitney U test. The Tukey HSD test confirmed all significant pairs.

6.5.2 *Personal Interpretation Response.* For all responses on signal clarity, the responses ranged from -3 (strongly disagree with the statement) to 3 (strongly agree with the statement).

Participants responded that they felt the most safe around the Sweeping set (median = 2.0, σ = 1.29) and the Car-like set (median = 2.0, σ = 0.94), followed by the Heartbeat set (median = 1.5, σ = 1.46) and lastly the Standard set (median = 1.0, σ = 1.54). Results are shown in Figure 17a. The light set did not significantly impact how safe participants felt. This was tested using the Kruskal-Wallis test (F(3,115) = 5.57, $p > 0.05$).



(a) "I feel **safe** around a robot with these lights."

(b) "I feel **comfortable** around a robot with these lights."

Fig. 17. Survey responses to "I feel **safe, comfortable** around a robot with these lights". Responses are on a seven point Likert scale from -3 (I strongly disagree with this statement) to 3 (I strongly agree with this statement). Shown as median, 25% quarantile, 75% quarantiles, and outliers by light pattern. Significance is shown as * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$.

Participants responded that they felt the most comfortable around the Sweeping set (median = 2.0, $\sigma = 1.58$) and the Car-like set (median = 2.0, $\sigma = 1.60$), followed by the Standard set (median = 1.5, $\sigma = 1.62$) and lastly the Heartbeat set (median = 1.0, $\sigma = 1.56$). Results are shown in Figure 17b. The light set did not significantly impact how comfortable participants felt. This was tested using the Kruskal-Wallis test (F(3,115) = 0.37, $p > 0.05$).

Participants responded that they felt the most distracted by the Sweeping set (median = 1.0, $\sigma = 1.50$) and the Car-like set (median = 1.0, $\sigma = 1.50$). Participants were the least distracted by the Standard set (median = -2.0, $\sigma = 1.42$), followed by the Heartbeat set (median = -1.5, $\sigma = 1.38$). LED light set (including the Standard) significantly impacted how distracted participants were (F(3,115) = 38.00, $p < 0.001$) using Kruskal-Wallis test. The Mann-Whitney U test was used to compare the custom light sets to the Standard light set, as seen in Figure 18. The Sweep set was rated significantly more distracting than the Standard set (U = 146.5, $p < 0.001$), as were the Car-like set (U = 129.5, $p < 0.001$) and less significantly the Heartbeat set (U = 320.5 $p = 0.024$). Significance between the custom light sets was tested using the Mann-Whitney U test. The Sweeping set was significantly more distracting than the Heartbeat set (U = 215.0, $p < 0.001$), as was the Car-like set (U = 189.5, $p < 0.001$). The Tukey HSD test confirmed all significant pairs. Results are shown in Figure 18.

6.5.3 Participant favorites. Overall, all the participants agreed that the custom light sets were more communicative than the Standard lights. No one chose the Standard lights alone as the favorite set they saw, as seen in Fig. 19. Some said they would have the Standard lights with a combination of the other patterns, but no one thought that the Standard ones were best. However, 16.7% of participants did cite the Standard set of lights as their least favorite, due to a lack of information being communicated by them.

Out of the three custom sets, Sweeping and Heartbeat were most often chosen as the favorite stand alone sets. This contradicts the original hypothesis that the Car-like lights would be the favorite. Conversely, 22.6% participants also chose Sweeping set as their least favorite. Heartbeat was also well liked by 25.8% participants because it was only a single color that could be seen from any angle. This contradicts the hypothesis that Heartbeat would be the least liked set overall. Some participants also said the Heartbeat lights felt calmer to be around than the other sets and liked that it was minimal while still conveying more information than the original lights. However,

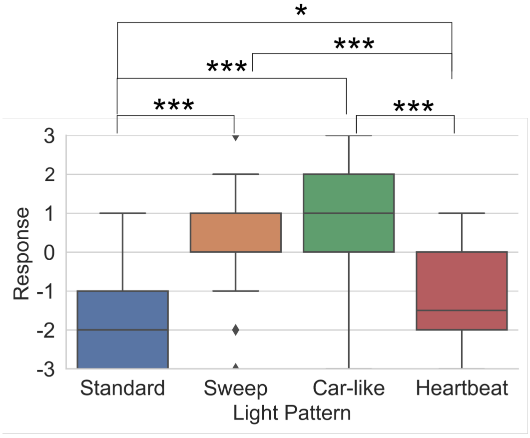


Fig. 18. Survey response to the statement: **I feel distracted by these lights**. Responses are on a seven point Likert scale from -3 (I strongly disagree with this statement) to 3 (I strongly agree with this statement). Shown as median, 25% quarantile, 75% quarantiles, and outliers for the four light patterns in response Significance is shown as * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$.

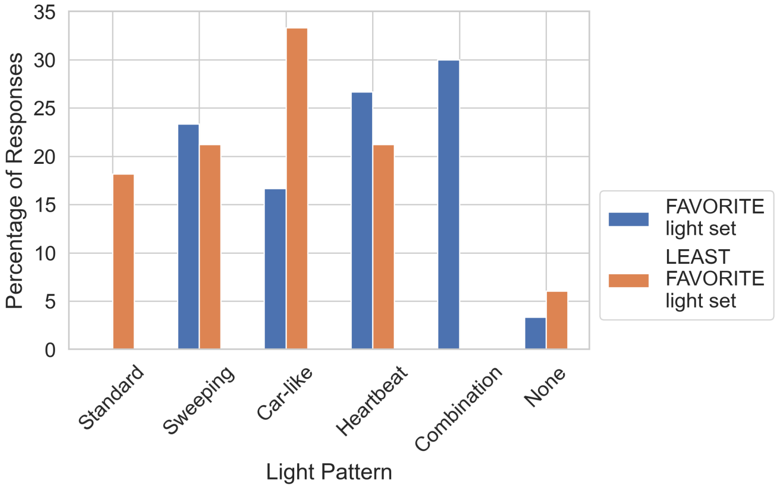


Fig. 19. Percentage of responses that voted for each light set as Favorite and Least Favorite.

using only one color and full bar pulsing had the drawback of looking similar to Standard lights for emergency stop and work-in-place for some participants. The single color also made it difficult for many participants to tell which side was solid and which was pulsing during the turns. Car-like was the least liked set of the three, with only 16.1% participants choosing it as their favorite, but 35.5% choosing it as their least favorite. This contradicts the hypothesis that Car-like would be the most liked set overall. People also generally stated that they felt safest around this light set due to familiarity, which supports the original hypothesis.

6.6 Evaluation between Employee Teams and Robot Familiarity

To ensure that there were no major biases between the participants who worked on the robots and their components and those who did not, pairwise Mann-Whitney U tests were run to assess if there was any significant difference between the two participant groups. For the most part, there was no significant differences between participants who worked on the robots and those who did not. The only exception was the Heartbeat light set, in which two states "turn" and "idle" were highly significantly different ($p < 0.005$ in both), for which participants who did not work on the robot were rated the Heartbeat lights with higher signal clarity than participants who work on the robots. Participants who did not work with the robot rated the signal clarity of "turn" state for Heartbeat as 1.69, whereas those who worked on the robot rated it as -0.06. Similarly for the "idle" state, participants who did not work on the robots rated the signal clarity as 1.54, and those who did rated it at 0.06.

The same tests were run between the responses of participants who were familiar with the robots and their lights and the participants who were not familiar with the robots and their lights. Like in comparing the employee teams, there was little significant difference in the signal clarity or participant interpretations between participants who were very familiar with the light pattern and those who were not familiar with the light patterns. The one light set with significant difference was signal clarity for the the Car-like light set at the "at goal" and "idle" states. In these two states, those who were familiar with the robot light patterns were more favorable of the Car-like light pattern than those who were not. For the "at goal" state, participants familiar with the robots had a significantly higher ($p < 0.05$) average signal clarity rating of 1.86, whereas participants who were not familiar had an average signal clarity rating of 0.81. Similarly for the "idle" state, participants with robot and light familiarity rated the Car-like lights significantly more favorably ($p < 0.05$) than those without familiarity with average signal clarity ratings of 1.64 and 0.75 respectively. One possible reasoning for this difference between the two populations is that those familiar with the lights likely know that the Standard lights do not signal for either "at goal" or "idle" states, causing them to comparatively view the Car-like lights as more clear. The Car-like lights are the most similar to the Standard lights out of the custom sets, which may have created an unconscious comparison between the two lights for those familiar with the Standard lights.

7 DISCUSSION

In this work, our research goals were to utilize the knowledge and experience of Clearpath employees in conjunction with knowledge from social robotics research to create legible, communicative lights for low degree of freedom factory robots. Using a company survey to integrate employee perspective into the design of the lights and a user study to evaluate our designed light sets, we found several insights for continuing work in communicative, legible lights for factory robots.

People making and working with robots have valuable insights about their communications: Using an online survey, we were able to collect concepts for styles of light designs from existing employees and create custom light sets based on these concepts. From this survey, we learned that in a factory setting, at least according this group, lights were the most desired channel of communication. This may be because they can be seen at a distance and do not interfere with the robot's current task. In contrast, the motion channels were seen as interrupting the robot task, and the sounds might not be heard in a loud factory environment. The Car-like lights were inspired by the company survey results and were based on the analogy of a car, with off-white headlights in the front, and dimmer red lights in the back in the "drive" state. For "blocked," "turning," and "at goal" states, the rear and headlights flashed in various colors. The Sweeping lights were based on the results of the company survey. The intention behind them was to have the lights move around the robot and converge to a particular point of interest, whether that be the goal or and obstacle. The colors for the states in the custom set were based on employee suggestions in the company survey. The Heartbeat "persona" was inspired by minimal light communication with colors, as seen in the work of Kim Baraka, which showed that red pulsing light the best combination for indicating a blocked state [3].

This participatory approach surfaced desired robot state communications that were needed, but did not previously exist: The original light set included "turning" lights, but we added "blocked," "at goal," and "idle." For example, many people mentioned the importance of a robot communicating when it is blocked so it can continue its operations. Additionally, participants in the study mentioned how they did not think about a "blocked" signal before, but once they saw it, they found the Standard lights lacking. Another direct request was an "idle" state so people, e.g., the test team at the company, could understand its operational status.

Familiarity of the robots, or working on them directly, may not predict how people view light communications: We ran comparisons between the responses of participants who work on the robots vs participants who do not and participants who are frequently interact with the robots and are familiar with the light signals vs participants who did not. Results showed that the team participants worked on and their familiarity with the robots generally did not affect how they saw the lights, both in signal clarity and personal response. This lack of connection between familiarity and signal clarity is useful, since the robots are used at many different factories and there is no one model for how often end users interact with the robots or how familiar they are with their communications.

In different situations, different styles of light communication are preferred: Sweeping had the most motion and was good for the "blocked" situation because it grabbed attention and showed the robot needed something to move to complete its task. Across all participants it was clear that the "blocked" state lights were the most clear in the Sweeping set. The "blocked" state of the robot is one that requires attention from a user, either because the user needs to move an object from the robot's path or the user needs to move themselves. Even though some participants found the the Sweeping lights overwhelming and distracting, most pointed out that was actually desirable when the robot was in "blocked" state. This finding supports the original hypothesis. Similarly, the for "at goal" state, the Heartbeat set was preferred, which also supported the original hypothesis. Participants cited that they liked that it was visible from any angle, which is ideal for directionless states like "at goal," but possibly not as useful for states where indicating directionality would be helpful, like "blocked" or "drive." Participants also liked the Sweeping set, again because it could be seen from all angles. The Standard lights and Car-like lights were the most desired for turning. They were very similar and were both familiar signals seen in driving. Again supporting the original hypothesis, Car-like was the favorite for "turning" state. Most participants stated that it looked like a turn signal, which was a familiar analogy to them. A signal was needed to indicate that the robot was idle, but the ones tested here were all too subtle to be fully communicative, especially from afar. Car-like was also chosen as the favorite for "idle" state, but all three custom sets have room for improvement. Merely dimming the lights to show the robot is idle may not be overt enough for it to be obvious to a user.

All of the new light sets outperformed the factory default: The addition of new states and employee-inspired designs caused user study participants to favor the custom light sets over the Standard lights. People said they liked the industry default lights until they saw the other light sets. After seeing the custom light sets, participants remarked that they thought the Standard lights did not communicate enough for most situations. However, Standard lights were non-distracting, likely because everyone is familiar with the industry default lights. All participants saw the Standard light set first, and during the first interview many participants seemed to like the Standard lights.

However, after they saw the first set of custom lights, many remarked in the second interview that they had not realized how lacking the Standard lights were.

8 CONCLUSION

In this work we set out to create more legible and communicative light patterns for factory robots. We used participatory design methods to create a survey to gain design insight from company employees. We then used this information, paired with prior work in social robotics to create three custom light sets for an OTTO 100 robot, evaluating these light sets against the original light implementation with a user study. Versions of these lights are deployed at over 30 sites on over 250 robots, which have completed over 150,000 hours of operation, travelling over 100,000 km in high-pressure factory environments.

Overall, our results showed that all three of our custom light sets out-performed the existing robot light sets, in part because they sought to clarify a greater number of robot operations. Rather than finding one new custom set to perform best overall, we instead found that different light sets were preferred for different contexts. For example, car analogies worked well when the robot was traversing spaces similar to on-road driving, where cars also effectively use their lights to communicate. However, in situations without a clear flow of traffic or that involved interaction, alternative light sets, such as the Sweeping or Heartbeat were often preferred. Heartbeat, a pulsing light, worked well for situations that were not urgent, as the lights were subtle and less distracting. On the other hand, situations that involve directionality, e.g., 'I'm trying to go there, please move out of my way," may instead benefit from sweeping.

Portions of the code we developed have already been deployed on robots operating in factories and warehouses across North America. Thus, the logical next step in this work would be to conduct further analyses of robot communications with factory workers during actual operations, which involves more challenging organization in terms of getting the data permissions across several locations, but would also assess how our results translate to a real world setting, i.e., integrated into a factory/warehouse workflow with all the typical distractions of a worker getting their job done. Such work could benefit from ethnographic approaches similar to our previous work results [10], to see what workers think of the new lights and how/when they are most effective.

This work also demonstrates the benefits of integrating domain expert perspectives, e.g., the knowledge and experience of the robot development team, into the application-centric robot communication design. For example, some participants expressed preferences for different light sets based on their own aesthetics, saying that the sweeping lights were fun to watch, whereas others felt overwhelmed, saying it was like “a 90s rave.” Thus expanding the robot's perception capabilities to be able to detect particular individuals in a factory may help in customizing machine communications, or using such communications for social signalling, for example, when arriving to hand off a part or object to a person. Communication design decisions may contribute to the branding or worker experience of a particular workplace.

Finally, the insights gained from the company survey relative to sound and motion communications, joined with research results from social robotics could follow a similar process for continued refinement of the OTTO robots multi-modal communication with people. This work can serve as a stepping stone and a process guide for designing effective communication with the insight of employees. While our final technology implementation focused on light patterns, we hope this work will act as the "Hello World" for a range of multi-modal robot communications designed with the needs and opportunities of factory settings in mind.

ACKNOWLEDGMENTS

Thank you to all the employees at Clearpath who filled out surveys and participated in user studies. Special thanks to the Testing Team for allowing us to borrow OTTO 100 #41 for an extended period, and to the Autonomy Team for helping debug and create the software for custom lights.

REFERENCES

[1] Abhijeet Agnihotri and Heather Knight. 2019. Persuasive ChairBots: A (Mostly) Robot-Recruited Experiment. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–7.

[2] Rachid Alami, Alin Albu-Schäffer, Antonio Bicchi, Rainer Bischoff, Raja Chatila, Alessandro De Luca, Agostino De Santis, Georges Giralt, Jérémie Guiochet, Gerd Hirzinger, et al. 2006. Safe and dependable physical human-robot interaction in anthropic domains: State of the art and challenges. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 1–16.

[3] Kim Baraka, Ana Paiva, and Manuela Veloso. 2016. Expressive lights for revealing mobile service robot state. In *Robot 2015: Second Iberian Robotics Conference*. Springer, 107–119.

[4] Kim Baraka, Stephanie Rosenthal, and Manuela Veloso. 2016. Enhancing human understanding of a mobile robot's state and actions using expressive lights. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 652–657.

[5] Kim Baraka and Manuela M Veloso. 2018. Mobile service robot state revealing through expressive lights: formalism, design, and evaluation. *International Journal of Social Robotics* 10, 1 (2018), 65–92.

[6] Andrea Bauer, Dirk Wollherr, and Martin Buss. 2008. Human–robot collaboration: a survey. *International Journal of Humanoid Robotics* 5, 01 (2008), 47–66.

[7] Regina Bernhaupt, Marianna Obrist, Astrid Weiss, Elke Beck, and Manfred Tscheligi. 2008. Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. *Computers in Entertainment (CIE)* 6, 1 (2008), 1–23.

[8] Baptiste Busch, Jonathan Grizou, Manuel Lopes, and Freek Stulp. 2017. Learning legible motion from human–robot interactions. *International Journal of Social Robotics* 9, 5 (2017), 765–779.

[9] Kathy Charmaz and Linda Liska Belgrave. 2007. Grounded theory. *The Blackwell encyclopedia of sociology* (2007).

[10] Bohkyung Chun and Heather Knight. 2020. The Robot Makers: An Ethnography of Anthropomorphism at a Robotics Company. *ACM Transactions on Human-Robot Interaction (THRI)* 9, 3 (2020), 1–36.

[11] Anca D Dragan, Shira Bauman, Jodi Forlizzi, and Siddhartha S Srinivasa. 2015. Effects of robot motion on human-robot collaboration. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 51–58.

[12] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 301–308.

[13] Brian R Duffy. 2003. Anthropomorphism and the social robot. *Robotics and autonomous systems* 42, 3-4 (2003), 177–190.

[14] Abrar Fallatah, Bohkyung Chun, Sogol Balali, and Heather Knight. 2020. " Would You Please Buy Me a Coffee?" How MicroculturesImpact People's Helpful Actions Toward Robots. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 939–950.

[15] Jaap Ham, René Bokhorst, Raymond Cuijpers, David van der Pol, and John-John Cabibihan. 2011. Making robots persuasive: the influence of combining persuasive strategies (gazing and gestures) by a storytelling robot on its persuasive power. In *International conference on social robotics*. Springer, 71–83.

[16] Chris Harrison, John Horstman, Gary Hsieh, and Scott Hudson. 2012. Unlocking the expressivity of point lights. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1683–1692.

[17] Bradley Hayes and Julie A Shah. 2017. Interpretable models for fast activity recognition and anomaly explanation during collaborative robotics tasks. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 6586–6593.

[18] Clint Heyer. 2010. Human-robot interaction and future industrial robotics applications. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 4749–4754.

[19] Guy Hoffman and Cynthia Breazeal. 2007. Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. ACM, 1–8.

[20] Jong-Hann Jean, Chen-Fu Wei, Zheng-Wei Lin, and Kuang-Yow Lian. 2012. Development of an office delivery robot with multimodal human-robot interactions. In *2012 Proceedings of SICE Annual Conference (SICE)*. IEEE, 1564–1567.

[21] Heather Knight, Timothy Lee, Brittany Hallawell, and Wendy Ju. 2017. I get it already! the influence of ChairBot motion gestures on bystander response. In *Robot and Human Interactive Communication (RO-MAN), 2017 26th IEEE International Symposium on*. IEEE, 443–448.

[22] Heather Knight, Ravenna Thielstrom, and Reid Simmons. 2016. Expressive path shape: simple motion features that illustrate a robots attitude toward its goal. In *2016 IEEE International Conference on Intelligent Robots and Systems (IROS). IEEE.*

[23] Heather Knight, Manuela Veloso, and Reid Simmons. 2015. Taking candy from a robot: Speed features and candy accessibility predict human response. In *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on.* IEEE, 355–362.

[24] Jun-tao Li and Hong-jian Liu. 2016. Design optimization of amazon robotics. *Automation, Control and Intelligent Systems* 4, 2 (2016), 48–52.

[25] Christina Lichtenthäler and Alexandra Kirsch. 2016. Legibility of robot behavior: a literature review. (2016).

[26] Diana Löffler, Nina Schmidt, and Robert Tscharn. 2018. Multimodal expression of artificial emotion in social robots using color, motion and sound. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction.* 334–343.

[27] Shaoping Lu, Chen Xu, Ray Y Zhong, and Lihui Wang. 2017. A RFID-enabled positioning system in automated guided vehicle for smart factories. *Journal of Manufacturing Systems* 44 (2017), 179–190.

[28] Nobuyoshi Matsumoto, Hiroyuki Fujii, and Michio Okada. 2006. Minimal design for human–agent communication. *Artificial Life and Robotics* 10, 1 (2006), 49–54.

[29] George Michalos, Sotiris Makris, Jason Spiliotopoulos, Ioannis Misios, Panagiota Tsarouchi, and George Chryssolouris. 2014. ROBO-PARTNER: Seamless human-robot cooperation for intelligent, flexible and safe operations in the assembly factories of the future. *Procedia CIRP* 23 (2014), 71–76.

[30] Dylan Moore, Rebecca Currano, and David Sirkin. 2020. Sound Decisions: How Synthetic Motor Sounds Improve Autonomous Vehicle-Pedestrian Interactions. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.* 94–103.

[31] Dylan Moore, Nikolas Martelaro, Wendy Ju, and Hamish Tennent. 2017. Making noise intentional: A study of servo sound perception. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI.* IEEE, 12–21.

[32] Michael J Muller and Sarah Kuhn. 1993. Participatory design. *Commun. ACM* 36, 6 (1993), 24–28.

[33] Bilge Mutlu and Jodi Forlizzi. 2008. Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction. In *2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI).* IEEE, 287–294.

[34] Sehoon Oh, Hanseung Woo, and Kyoungchul Kong. 2014. Frequency-shaped impedance control for safe human–robot interaction in reference tracking application. *IEEE/ASME Transactions On Mechatronics* 19, 6 (2014), 1907–1916.

[35] Sebastian Osswald, Roland Buchner, Astrid Weiss, and Manfred Tscheligi. 2012. Using participatory design to investigate technology usage in the cleanroom of a semiconductor factory. In *Workshop The Message in the Bottle: Best Practices for Transferring the Knowledge from Qualitative User Studies at DIS*, Vol. 12.

[36] Emma J Rose and Elin A Björling. 2017. Designing for engagement: using participatory design to develop a social robot to measure teen stress. In *Proceedings of the 35th ACM International Conference on the Design of Communication.* 1–10.

[37] Selma Šabanović, Wan-Ling Chang, Casey C Bennett, Jennifer A Piatt, and David Hakken. 2015. A robot of my own: participatory design of socially assistive robots for independently living older adults diagnosed with depression. In *International conference on human aspects of it for the aged population.* Springer, 104–114.

[38] Shane Saunderson and Goldie Nejat. 2019. How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics* 11, 4 (2019), 575–608.

[39] Douglas Schuler and Aki Namioka. 1993. *Participatory design: Principles and practices.* CRC Press.

[40] Sichao Song and Seiji Yamada. 2017. Expressing emotions through color, sound, and vibration with an appearance-constrained social robot. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI.* IEEE, 2–11.

[41] Sichao Song and Seiji Yamada. 2018. Bioluminescence-inspired human-robot interaction: designing expressive lights that affect human's willingness to interact with a robot. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction.* 224–232.

[42] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2015. Communicating directionality in flying robots. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* IEEE, 19–26.

[43] Leila Takayama and Caroline Pantofaru. 2009. Influences on proxemic behaviors in human-robot interaction. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems.* IEEE, 5495–5502.

[44] Astrid Weiss, Roland Buchner, Manfred Tscheligi, and Hanspeter Fischer. 2011. Exploring human-robot cooperation possibilities for semiconductor manufacturing. In *2011 International Conference on Collaboration Technologies and Systems (CTS).* IEEE, 173–177.

[45] Daniela Wurhofer, Verena Fuchsberger, Thomas Meneweger, Christiane Moser, and Manfred Tscheligi. 2015. Insights from user experience research in the factory: What to consider in interaction design. In *IFIP Working Conference on Human Work Interaction Design.* Springer, 39–56.