



DEGREE PROJECT IN COMPUTER SCIENCE AND ENGINEERING,
SECOND CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2016

Towards Enhancing Human-robot Communication for Industrial Robots: A Study in Facial Expressions

LAN WANG

Towards Enhancing Human-robot Communication
for Industrial Robots:
A Study in Facial Expressions

Mot Förbättra Människa-robot Kommunikation
för Industrirobotar :
En studie i ansiktsuttryck



**KTH Computer Science
and Communication**

Lan Wang
lanwa@kth.se

Supervisor: Vygandas Vegas Simbelis
Examiner: Mario Romero Vega
Principal: ABB Corporate Research

*Degree Project in Computer Science and Engineering
School of Computer Science and Communication*

KTH CSC
SE-100 44 Stockholm, Sweden

2016-07-06

ABSTRACT

Collaborative robots are becoming more commonplace within factories to work alongside their human counterparts. With this newfound perspective towards robots being seen as collaborative partners comes the question of how interacting with these machines will change. This thesis therefore focuses on investigating the connection between facial expression communication in industrial robots and users' perceptions. Experiments were conducted to investigate the relationship between users' perceptions towards both existing facial expressions of the Baxter robot (an industrial robot by Rethink Robotics) and redesigned versions of these facial expressions. Findings reveal that the redesigned facial expressions provide a better match to users' expectations. In addition, insights into improving the expressive communication between humans and robots are discussed, including the need for additional solutions which can complement the facial expressions displayed by providing more detailed information as needed. The last section of this thesis presents future research directions towards building a more intuitive and user-friendly human-robot cooperation space for future industrial robots.

ABSTRAKT

Delade robotar blir allt vanligare inom fabriker för att arbeta tillsammans med sina mänskliga motsvarigheter. Med denna nyfunna perspektiv mot robotar ses som samarbetspartners kommer frågan om hur interagerar med dessa maskiner kommer att förändras. Denna avhandling fokuserar därför på att undersöka sambandet mellan ansiktsuttryck kommunikation i industrirobotar och användarnas uppfattningar. Experiment utfördes för att undersöka sambandet mellan användarnas uppfattning mot både befintliga ansiktsuttryck av Baxter robot (en industrirobot av Rethink Robotics) och omgjorda versioner av dessa ansiktsuttryck. Fynden visar att de omgjorda ansiktsuttryck ger en bättre matchning till användarnas förväntningar. Dessutom är insikter förbättra uttrycks kommunikationen mellan människor och robotar diskuteras, bland annat behovet av ytterligare lösningar som kan komplettera de ansiktsuttryck som visas genom att ge mer detaljerad information om det behövs. Den sista delen av denna avhandling presenterar framtida forskningsinriktningar för att bygga ett mer intuitivt och användarvänligt människa-robot samarbets utrymme för framtida industrirobotar.

Towards Enhancing Human-robot Communication for Industrial Robots: A Study in Facial Expressions

Lan Wang

Royal Institute of Technology

Stockholm, Sweden

lanwa@kth.se

ABSTRACT

Collaborative robots are becoming more commonplace within factories to work alongside their human counterparts. With this newfound perspective towards robots being seen as collaborative partners comes the question of how interacting with these machines will change. This thesis therefore focuses on investigating the connection between facial expression communication in industrial robots and users' perceptions. Experiments were conducted to investigate the relationship between users' perceptions towards both existing facial expressions of the Baxter robot (an industrial robot by Rethink Robotics) and redesigned versions of these facial expressions. Findings reveal that the redesigned facial expressions provide a better match to users' expectations. In addition, insights into improving the expressive communication between humans and robots are discussed, including the need for additional solutions which can complement the facial expressions displayed by providing more detailed information as needed. The last section of this thesis presents future research directions towards building a more intuitive and user-friendly human-robot cooperation space for future industrial robots.

Author Keywords

Industrial Robotics; Human-robot Collaboration; Facial-expression communication; Anthropomorphism; Nonverbal Communication.

1. INTRODUCTION

Factories are well accustomed to industrial robots. They used to be described as big, strong and robust devices, surrounded by fences to warn people how dangerous they were. Currently focus is placed on human-robot interaction, such as reducing harm to humans in collisions. The rise of automation has empowered new capacities which previously only humans did [1], such as pick and place tasks, and assembly operations. These techniques can be

seen in many industrial robots nowadays, such as the YUMI robot from ABB, and the IIWA robot from KUKA [22].

Meanwhile, screen-based teleoperation has been concerned with using mobile robots to perform maintenance and manipulation in factories, with its interfaces using multiple 2D views or even 3D view [1]. Prominently, these techniques are being used to facilitate a greater degree of adaptive automation.

While robots were initially used in repetitive and simple tasks, they are becoming involved in increasingly more complex and less structured tasks, including collaborating with people to complete tasks together. As described in Robert's survey [22], the Kuka robots without protective barriers have been used to pick and pass the components to assembly workers at Audi's Ingolstadt production facility in 2015. However, there is still minimal feedback in terms of signals to the operator, with the robot's internal state essentially invisible [4]. When a robot performs a task, the operator may easily get confused as to why a request could not be executed. For instance, when a robot stops its consecutive actions unexpectedly, it may not be obvious to the operator that the executive module broke down or that the object could not be identified. Thereby eliciting an expressive connection between robots and operators shows potential in the industrial area; thus, making the robot's intentions and states more understandable and predictable to humans.

This master thesis is conducted at and together with ABB Corporate Research [31]. The principal interest is to investigate the effectiveness of current facial-expression communication used in human-robot interaction for collaborative robots [22], in order to explore the necessity and future direction of designing industrial robots with more social components as opposed to being engineered with strictly functional affordances, as they are today. The target market is defined as the future factory, which corresponds to the novice user without prior experience of

robotic programming and production. Only nonverbal expressions for industrial robots were studied in this thesis. The capabilities of robots to sense the environment and interpret the operators' behaviours are not involved, such as the speech recognition, human facial expression detection, etc.

1.1 Research Question

The aim of this thesis is to explore the value and possibility of adding facial expressions to industrial robots, through evaluating and redesigning existing facial expressions of the Baxter robot. Thus, the research question is defined as follow:

Can people understand industrial robot behaviours better through the redesigned Baxter-like facial expressions?

In order to answer this question comprehensively, the more specific sub-questions are established in consecutive order of execution:

1. Does the existing set of facial expressions for current industrial robot solutions (i.e. the Baxter robot behaviours) map accurately or not to users' understanding?
2. If some of the gestures are not understood correctly, how can that be complemented or redesigned to match user's understanding better?
3. Does the redesigned version of these facial expressions map to users' understanding better than the Baxter-like facial expressions?

1.2 Innovation

The social expressive component is supposed to elicit an emotional connection between operators and industrial robots, making operators understand the robotic intents and "what they are thinking", thus, possibly easing the cognitive workload and reduce training cost of the operator. After the preliminary research on this direction, it is believe that the human-robot communication based on facial expressions is an area which requires continued research.

2. BACKGROUND

This chapter elaborates the findings during the literature review towards the fields of human-robot collaboration, anthropomorphism and nonverbal communication. Related works are included in the section of each field.

2.1 Human-robot Collaboration

Industrial robots are mostly engineered with specific affordances for relatively controlled environments. Although they are increasingly intelligent and powerful, the cost-intensive issue on programming by experts does still matter. Recently there has been a tendency in the industrial area to extend the focus from separated autonomous robots

to cooperative robots [5] that can operate alongside humans in a shared workspace without conventional protections [22] (i.e.fencing), and with which operators can communicate and interact. As robots and people begin to coexist and collaborate, natural interaction between human and robot which resembles human-human interaction is becoming more and more important. The industrial direction changes from functional robots to "buddy" robots, which have been considered with a shared workspace and a "shoulder-to-shoulder" collaboration [13].

Human-robot collaboration has been researched in numerous previous works. Guy Hoffman and Cynthia Breazeal have investigated turn-taking and joint plans in the context of verbal and nonverbal dialog, based on which a theoretical framework was established and applied to their humanoid robot Robonaut [8]. This was envisioned to work with human astronauts. Afterwards, they evaluated the collaborative fluency in human-robot shared-location teamwork with a set of metrics [9]. In their later research, a cognitive architecture was built based on the euro-psychological principles, initially from the study measuring team efficiency and fluency. The embodied result was a robotic desk lamp AUR, performing in a human-robot collaborative task [7]. Other than the state-of-art robots proposed for social communications, there are lots of studies towards the collaborative robots in the industrial field. Early work has studied a robotic arm helping people with an assembly task through the vision-recognition systems (Kimura et al. 1999), which constructed a framework for the human-robot cooperation at an initial stage [23]. Mechanical coordination and safety considerations are consistent with recent laboratory research in the field of shared-location human-robot collaboration. However, as mentioned in the study of Fairul et al. [13], physical safety has to be complemented by "mental safety", for instance, by the consciousness of robot motion.

However, all of the above related works are limited to laboratory research [12]. So far, there are limited practical applications upon human-robot teamwork in the industrial field. One notable exception is the robot Baxter by Rodney Brooks [6]. Baxter is defined as a unique industrial robot with "common sense", combining appearance elements of an industrial and a humanoid robot. In addition, it is designed in the context of a shared workspace and adopted with the kinesthetic teaching approach. However, the research support or theoretical argument for the usability and effectiveness of its facial expressions can not be found in the academia during the literature searching. As a first step towards understanding the effectiveness of this type of approach, an investigation into whether the set of facial expressions presented by the Baxter robot can be deemed feasible is needed.

2.2 Anthropomorphism

Within social robotics [35] and service robotics (a type of robot that performs services useful to humans excluding industrial automation application [29]), the use of anthropomorphism design elements can be seen, such as in robot head design.

Mainly three types of humanoid robot heads can be seen in the field of social robotics, among which the number of facial expressions and the influences on people's perception differ [30]. The first type is interactive robot heads, based on speakers and LEDs, which have a larger usage in social and service robotics. For instance, NAO robot [30] uses an interactive head with LED "eyes" and a speaker to express emotion. It is featured with simplicity, lighter weight, and the capability of displaying a wide range of colours to make eye expressions. While this brings an obvious problem of the interpretation of the facial expressions. The kinematic robot head, the second type, is built with various movable parts to make facial expressions and body gestures in a primitive way, which can be seen in the existing works, such as the robot Kismet [34] from MIT. They are engineered to provide a communication significantly closer to the natural communication mode between humans, for one thing, which leads to an increased engagement with the human, for another, also along with the shortage of complexity. The use of animatronic head with flexible skin, such as the Alice robot and the Albert HUBO [30], can be regarded as the third type of robot head design. These realistic robots express emotions mainly through mimicking human faces and the movements. It is the most precise technique to match the robot facial expressions to the words spoken, while currently it shows that they make people feel uncomfortable.

In personal companion robotics, the wide-spread adoption of zoomorphic or abstract appearance has been combined with the anthropomorphic behaviours, as exemplified by the robot Jibo. Jibo's design follows the "Kindenschema" [32] principle ("baby schema"), using the facial expressions on a screen and body gestures to engage with its users.

With the tablet computer advancing in sensor-rich units and extended computation and communication capabilities, it shows high potential in human-robot interaction [14]. Lots of works have employed a tablet head mounted on the robot or even as the robot itself. Such as the MIT robot Tega, the mini phone-based robot Romo, and the android-based robot ChibiFace [14]. Since tablets can be programmed more easily to add various functions to the robot head such as vision, hearing and speaking, and also they are more portable and customizable, the tablet-based robot head may be widely used in the near future. We regard this type of robot head as the study object in this thesis, in order to preliminarily explore the possibilities of designing industrial robots with more social components.

On one hand, attaching human traits, emotions and intention to non-human entities might be easier to elicit people's empathy and lead to more effective and realistic communication between robots and humans. On the other hand, the "Uncanny Valley" [26] exists, where the "almost" anthropomorphism would impact human's familiarity of the robot. Hence an important aspect of the future robot design is to balance the humanoid robots with the cartoon-like design elements.

2.3 Nonverbal Communication

Humans show diverse emotions through different kinds of channels. Each emotion consists in specific contexts and influences the behaviours and expressions of the human as well as that of other people next to the human. Emotions enhance natural human-human communication, help people understand each other's inner state and then may help decide appropriate responses to that person. The effect of a study on human-computer interaction shows that humans tend to react to computers in the same way that they do with other humans [15], especially in some long-term "relationship" with computers. Thus, researchers believe this can also apply to human-robot interaction.

In fact, the ability to express emotions has been described as one of the indicators of socially interactive robots [2], which utilise verbal or nonverbal channels to express their emotions. Verbal communication (e.g. speech) could aim primarily at passing text messages, while nonverbal communication (e.g. body movement, posture) is better for conveying spatial information. Admittedly, using speech to express "what they are thinking" is intuitive for many people, and speech has also been examined for multitudes of social or service robots, including the robot head Furhat from KTH, the companion robots Buddy and Pepper. Yet speech is not always sufficient or straightforward for passing lower-level details [3], which can be remedied by nonverbal communications.

Nonverbal communication is defined as sending and receiving unspoken clues. It covers several domains [27], including Kinesics (body language), Proxemics (distance), Paralanguage (e.g., voice quality), Haptics (touch), Oculistics (e.g. eye contact), etc. Haptics and Proxemics have been applied in extensive robotic research, and also can be seen in industrial robotics. Such as the tangible control buttons on the body of industrial robots, and the considering of safe operating area in manufacturing [1]. A wide range of studies following Oculistics and Kinesics is focused on the facial display of robots, which has been the most common mechanism for expressing non-verbal affects. This research can be seen in the aforementioned humanoid robot heads and tablet-based robot heads. Some robots have used other nonverbal cues, like body movement, orientation, colour, and sound (one application of Paralanguage). These can be seen in the survey by Bethel

and Murphy [10], which studied affective communication in appearance-constrained robots for naturalistic human-robot interactions. The result suggests to use robot orientation and sound to show attentiveness and caring for humans in an intimate proximity zone, while combining the body movement and posture in the personal and social zones. Other similar work has been conducted in different areas, such as the robot lamp in the movie *Luxo Jr.* (Pixar 1986), and a peripheral robotic companion Kip1 [25] for promoting peaceful conversation between humans.

A study from Carnegie Mellon University [11] investigated how the robot's emotional expressions influenced people's perceptions and behaviours, testing on the Roboceptionist with an LCD head displaying facial expressions. The result showed that simply changing the expressive emotion of a robot had a strong effect on people's behaviour and interactions with it. Another experiment by Rahman et al. [12] employed a robot Baxter to collaborate with humans for the assembly task in manufacturing, with the result showing that the robot displaying static emotion produced better human-robot interaction and assembly performance than that the robot produced with no emotion. And they believe that dynamically changing the robot's emotion based on task situations may enhance more of the human-robot collaboration.

It is noted earlier that visual cues have a stronger response than audio cues. As a first step towards studying how increasing expressions employed on robots affect user's understanding, this thesis focus on nonverbal communications which visualising robot's internal state transparently in case information conveyed through audio is ambiguous or inefficient. In other words, using nonverbal channel to disclosure robot emotion with facial expressions is opted to be studied in this thesis, instead of the verbal expression.

3. THEORIES

This chapter describes one set of principles and two physiological theories used in the later stage of redesigning the robot facial expression. The redesign section was inspired by practices and principles of animation from Disney, combining with some somatological theories, and applied them in enhancing humanoid expressions of robots to be more understandable for humans.

3.1 Disney's Twelve Principles of Animation

Johnston and Thomas proposed the Disney's Twelve Principles of Animation in 1995 [16], which is not only followed by Disney's animators but also affecting the majority of animated films nowadays. Animation artists try to create believable emotional characters and depict the illusion of life, as claimed by Van Breemen in 2003 [17], which is missed both in early day's animations and the user-interface robots. Based on this, he proposed to apply the

Twelve Principles of Animation to robots, especially the creation of robot animated expressions.

The Twelve Principles of Animation includes Squash and stretch, Anticipation, Staging, Straight ahead and pose-to-pose, Follow-through and overlapping action, Slow in and slow out, Arcs, Secondary action, Timing, Exaggeration, Solid drawing, and Appeal [16]. Most of the principles above can be applied to robotics [18]. For example, Anticipating actions contributes to guessing what a character is going to do next, which can be used for users to better interpret the robot. Staging suggests adding multi-modal interactions to robots like sound or lights, to make the expressive intention clearer to users. The principle of Timing described that the same movement can have different emotions depending on the timing used, which can be also applied to robot's expressions. In addition, Exaggeration is suitable for robot's gestures and expressions to emphasise the actions and make them more noticeable. These principles and practices from Disney's animations were used as the inspiration during the redesign process.

3.2 Wide-eyed theory

Generally humans will become wide-eyed when they are scared or stimulated. Daniel Lee and Adam Anderson [21] did research on the reason of this physiological phenomenon, finding that the wide-eyed expressions were useful as raw physical signals for not only providing the person who make the expressions a wider visual view, but also helping to send a clearer gaze signal telling observers to "look there". The wide-eyed theory can potentially be applied to robots that make expressions while getting stimulated, in order to catch users' attention quickly and help them locate the threats. This physiological theory was applied in the redesign of the *surprised* expression.

3.3 Pupillary response theory

Pupillary response is one kind of physiological response that brings size variations of the pupil. The two types of pupillary response include a constriction response [19] of narrowing the pupil and a dilation response with widening the pupil. In addition to drugs, the response of pupil dilation can be aroused in low light conditions to let more light into the eyes. As studied by Eckhard and James, the pupil dilation may indicate interest in the subject of attention [20]. This theory was utilised on the *concentrating* behaviour of the robot, which will be elaborated later in the redesign chapter.

4. METHODOLOGY

This thesis research started with literature study composed of the studies on four fields — industrial robotics, human-robot collaboration, anthropomorphism, and nonverbal communication. This gives a broader view upon the current state of industrial robotics and human-robot communication, furthermore, to narrow down the research scope and establish the research question. Besides, relevant theories towards animation and physiology were also researched in the whole study period, for the furtherance of the redesign phase. The articles and research papers are gathered through online databases like Google Scholar, and the useful findings are as the References at end.

In order to answer the research questions, a series of laboratory experiments were designed and executed thereupon. The whole process consists of the first experiment (Experiment I) investigating the existing set of facial expressions for industrial robots, a redesign phase based on results from the first experiment, followed by the second experiment (Experiment II) evaluating the redesigned facial expressions. Data analysis and some insights came after each experiment, with Experiment I looking into the data of the original facial expressions and Experiment II comparing the data before and after redesign.

The following paragraphs describe in detail how the laboratory experiments were processed as a method to investigate the facial-expression based communication in human-robot collaboration.

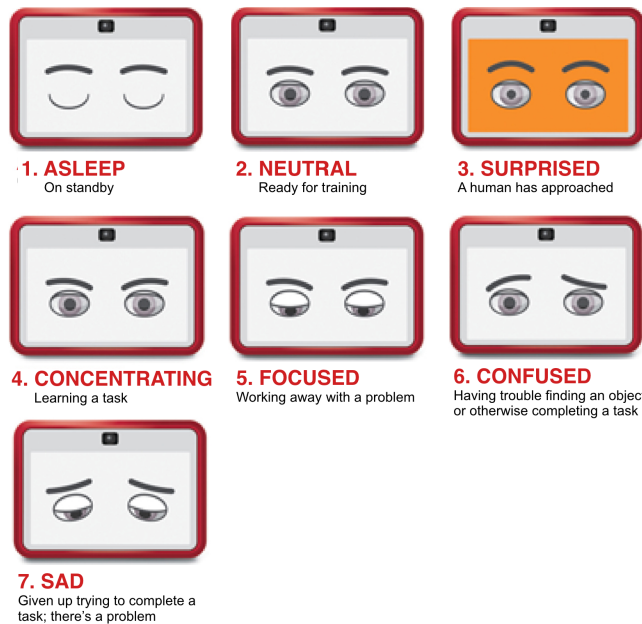


Figure 1: Official gesture setting of Baxter robot [33]

4.1 Experiment I

As a first step towards understanding the effectiveness of current facial expressions for industrial robots, an investigation into whether the set of behaviours presented by Baxter robot [6] can be deemed feasible is needed. Thus, we begin with investigating whether or not this set of facial expressions maps accurately to users' expectations and, furthermore, explore how can this work be extended. Such as through changing the facial expressions presented, increasing the level of detail in the design of these expressions, complementing these expressions with sound or actuators to increase head movement.

4.1.1 Experimental setup

We regard the current facial expressions for the Baxter robot solution as the baseline, as mentioned above, which is aimed to test the value and understandability of this set of facial expressions. There are seven behaviours set in the emotion of the Baxter robot, as seen in figure 1, including *neutral*, *asleep*, *concentrating*, *focused*, *surprised*, *confused* and *sad*. Each behaviour has its facial expression counterpart and a context of occurrence. For instance, when the robot is performing a task, it will show *concentrating* on the face, while it will have a *confused* look if the operator teaches it where to pick-up an object but forgets to show where to place the object. Based on these, a series of animations simulating the seven Baxter expressions were built in the software Adobe After Effect. Afterwards, the animated Baxter-like facial expressions were imported into an Android tablet, which was then placed on top of a dual arms industrial robot as a robotic head.

As another part of the evaluation system, a robot working scenario was constructed based on essential features of a industrial robot, including the human teaching and simple robot working task like pick and place. As figure 2 shows, the scenario consists of seven segments, with one robot behaviour embedded in each segment. Then a video was recorded in the laboratory following the whole scenario, and was processed in the software Adobe After Effect. Thus, a video-based scenario was prepared for presenting to the participants during the experiment, which is segmented into seven parts to map to the seven behaviours. There are several reasons as to why we choose video to present the robot working scenario to the participants, instead of using the real robot. Firstly, the form of video enables maintaining consistency between all the tests with different participants, and reducing the errors on the system, manual operation, etc. Secondly, a video-based test is beneficial to get a larger group of participants though bringing a portable laptop to all the participants, instead of arranging each test in the laboratory and setting schedules for each person to come.

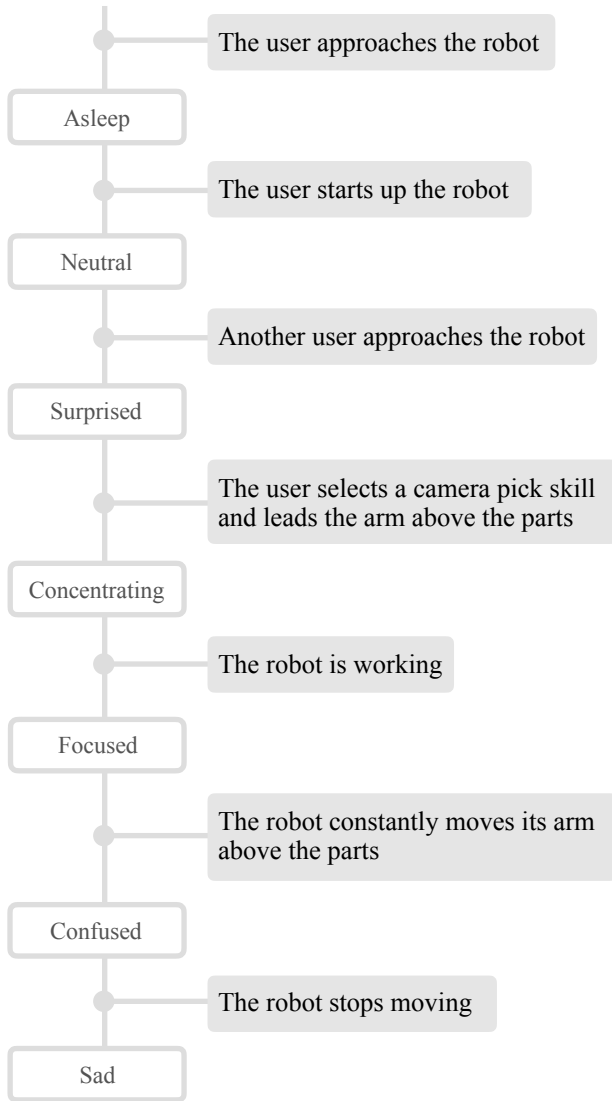


Figure 2: Robot working scenario

4.1.2 Data collection

Both quantitative data (from the Likert Scale questionnaires) and qualitative data (from the feedback during tests and the discussion session after tests) are collected during the experiment. A set of questionnaires were designed for collecting quantitative data during the experiment process, which includes 11 Likert scale built in one question for each segment (seen in Appendix 1), and one example at the beginning for describing how to rate. A behaviour list with 11 behaviour words corresponding to the 11 Likert scale was presented in each segment question, in which seven words described the seven original behaviours and another four words were “confusing words”

that we thought had high similarity with the meaning of some facial expressions. The four “confusing words” include *unsure*, *thinking*, *angry* and *expecting*. Under each segment, the participants were asked to indicate what they think the facial expression of the robot means, through rating one or more behaviours from the 11 behaviours list. The rating scale is from 0 (lowest) to 10 (highest), representing that the behaviour word is from mismatching to perfectly matching to the facial expression in the video. The scored range set as 10 points instead of 5 or 7 points (most of current Likert scale surveys choose the range of 0 to 5 or 0 to 7) is because a bigger range can capture more detailed variation towards people’s perception, getting relatively accurate results. During the process of filling out the questionnaires, participants were advised to talk about what they felt following the Think-aloud protocol [24] in order to capture the timely feedback on the presented facial expressions, while the tester would not talk or give any response to the participant to avoid bias. After completing the seven-segment video, there would be a free discussion session before the whole experiment done, during which participants were asked some relevant questions towards the value and effectiveness of the robots’ facial expressions. The questions were as follows:

1. Are the facial expressions easily for you to make sense? How?
2. Do you think this kind of robot facial expression is valuable for the human-robot communication? Why?
3. Do you think the information it provides is enough for you to understand the robot’s situation?
4. If not, what else components do you think could be added potentially to improve the human-robot communication?

The whole testing process was video-recorded (the testing environment is shown in Figure 3) so that the conversations and information gathered could later be traced back and reviewed if needed. Since the discussions following the experiments consisted of semi-structured interview sessions, this provided participants with an approach to more openly discuss the experiments, thus enabling valuable additional insights to be collected. The tests were around 15 minutes each, including 10 minutes’ video-based survey and 5 minutes’ discussion.

4.1.3 Participants

A total of 40 adults of mixed gender (21 male, 19 female), age 22 to 48 years, completed the laboratory experiments during this thesis research, with 20 adults attending each experiment respectively. All the participants are employees in ABB Corporate Research, without any vision, hearing or learning disability. They have a broad mix of background, such as design, engineering, mechanics, marketing,

accounting, logistics, etc. Meanwhile, as mentioned about the target user group, the participants have no relevant experiences in robotic programming and production, and also have not seen the Baxter robot before. The detailed participant population is shown in Appendix 2.

The participants selection followed the “samples of convenience” approach [28], in order to reach people easily. A series of invitations were sent out to people who fit for the requirements, which resulted in 40 participants with different ages, gender and backgrounds selected and divided into two groups to attend Experiment I and Experiment II respectively. The reasons for choosing this number of samples include: it can provide enough data to get relatively reliable results; the number was easily accessible in a short time frame.



Figure 3: Participants were watching the video and completing the questionnaires in the testing room

4.2 Data-based Redesign

This phase starts following the Experiment I, which achieved both quantitative and qualitative data towards the effectiveness of the Baxter-like facial expressions. Based on the results of Experiment I, some of the original facial expressions were considered to be ambiguous or misinterpreted. Hence these facial expressions were then redesigned or adjusted depending on the data analysis, collected insights, aforementioned theories and iterative brainstormings. Since the behaviours and contexts were kept the same, the solutions for each facial expression were generated with sketches, aimed at invoking a better map between the facial expression and its counterpart behaviour and context. Afterwards, a round of brainstorming was set up to discuss and reach the best solution.

Implementation and prototyping based on the redesign sketches followed. That used the same approaches as with

the setup of Experiment I. The redesigned facial expression solutions were created in the software Sketch, then the animation of each facial expression was conducted in the software Adobe After Effects separately.

4.3 Experiment II

To investigate whether or not the redesigned facial expressions map better to user’s understanding, an essential step after redesign is to evaluate it. The basic setup of Experiment II is consistent with the setup of Experiment I: the evaluation system was built with the same robot and same Android tablet; a video was prepared under the same scenario and same laboratory environment. Other parts that are unchanging between the both experiments, contain the questionnaires, the testing procedure, the sampling approach and the size of user group. The only variable in this experiment was the appearances and animations of the facial expressions presented in the tablet face, which kept the same graphic feeling but redesigned in different expressive way.

Another 20 participants who conformed to the speciality of target users and had not seen the Baxter-like robot interface before, were selected before running the Experiment II. Data was collected during the whole experiment including both quantitative data from the Likert Scale questionnaires, and qualitative data from the follow-up discussion and the recorded video.

5. RESULTS OF EXPERIMENT I

The Experiment I was executed over three days in March 2016 by 20 participants (named from P1 to P20), from which we got 20 questionnaire results and many insights in the discussion section. The questionnaire results provided quantitative data towards participants’ understanding on the facial expressions. The data were analysed by inspecting the number and distribution of choices, calculating the mean value of scores, while combining with the verbal feedback using Think-aloud protocol during experiments. Figure 4 shows the data analysis of Experiment I by using a confusion matrix [36].

As shown in Figure 4, most people (85%) mapped the *asleep* facial expression correctly, rating it with the highest mean value of 9.41. Meanwhile, 11 out of 20 people (55%) chose it as the *neutral* behaviour which got a mean value of 7.27, only second to the *asleep* one. As mentioned by the one participant (P3), “neutral is the behaviour that always exists when there is no obvious emotion” while another 5 people who chose the *neutral* behaviour also had similar comments. Two people thought the robot face looked like *asleep*, but also indicated *thinking* about something. It appears that the *asleep* feeling of this expression is not distinct enough, leading to speculation on other behaviours. Besides, P11 (who works at accounting and have no experience on robotics or cartoon-face) admitted it was hard

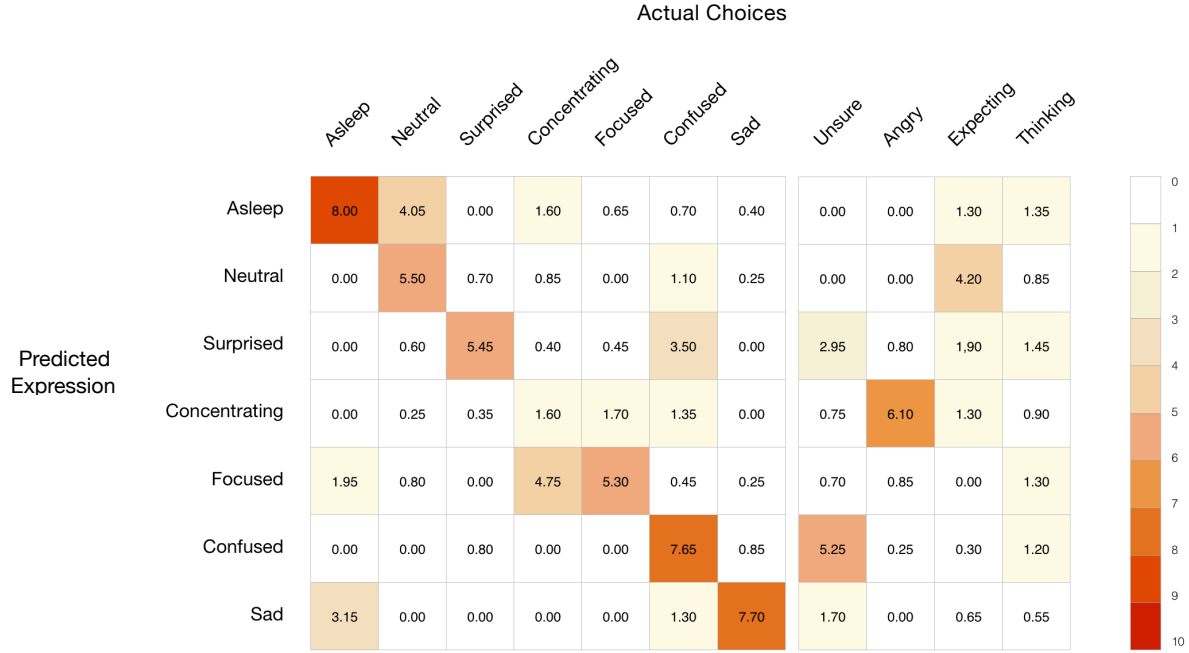


Figure 4: Data analysis of Experiment I.

The number shows the mean value of scores on each facial expression. Except for the seven original behaviours, the rated scores on other four confusing behaviours are shown on the right.

to understand what it was, especially the first time saw it. This seems due to the lack of relevant experience, which then may be the incentive of some ambiguous results. In addition, P13 considered the face as missing eyes, advising to add eyelashes.

For the *neutral* expression, the correct match got the highest mean value (8.46) with 13 people (65%) rating in a range of [6, 10]. Unexpectedly, there are 10 choices (50%) towards the *expecting* behaviour with a second high score of mean value (8.00). Mostly they thought the robot was “expecting what to do next” (P7). The reason appears that one facial expression (especially the one without much emotion) itself would have multiple interpretations, while other elements such as the context helps people differentiate the meanings. Thus, emphasising the context benefits the facial expressions understanding. Furthermore, it is interesting to see that the behaviours *focused*, *concentrating* and *thinking* got a few scores by the same group of people, who thought these 3 behaviours had “a very similar meaning”(P3). It seems that the meanings of these behaviours are ambiguous.

75% participants chose the *surprised* expression correctly with a high average score of 7.2, while 50% participants thought it as the *confused* behaviour marking with an average score of 7, and *unsure* also got 45% choices with an average score of 6.44. Some feedback towards why

choosing the *confused* and *unsure* behaviours includes, “the eyebrows raised mean the robot unsure what happened there” (P5) and “[I think] unsure and confused have the same meaning” (P11), which indicates the same problem above of the ambiguous behaviours. Additionally, the red colour on the whole face has various interpretations, such as “[the robot] is angry” (P3), “the robot shows unsafe” (P16), and “it is saying: ‘get away from me!’”(P5). Half of the participants mentioned the colour change is tricky, and one of them suggested, “the colour should not be on the whole face, but somewhere above the ‘head’”(P15). Furthermore, three people have the comments: “eyes should shift from one person to another during the facial expressions change” in order to make sense and notify the user.

The *concentrating* expression got the most inaccurate result compared to others: 16 out of 20 people (80%) regarded the facial expression as an *angry* look with the average score of 7.63, while the correct match — *concentrating* behaviour, only had 6 choices (30%) with an average result of 5.33. Meanwhile, some scores were scattered about among other options, including *focused*, *unsure*, *thinking*, *confused* and *expecting*. Considering both the scores data and oral feedback, it seems the in-tilted eyebrows producing an *angry* feeling that results in the incorrect choices. Another main reason mentioned is “the lack of eye-following” (P5). It is worth noting that 4 people were not sure what

happened here since “the facial expression changed too slightly to catch attention”(P11).

From Figure 4 we can see that the *focused* expression gets 13 correct choices (65%) as much as the *concentrating* option, while the mean value of *focused* is 8.23, larger than that of the *concentrating* one (7.38). This could be because the two words, *focused* and *concentrating*, seem “not much differences”(P3). Following that, 5 people (25%) thought it an *angry* face as it met the same eyebrows problem with the facial expression in segment 4. It was thought as an *asleep* look with a high average score of 7.8, in consequence of “the eyelids almost closed” with the suggestion to “make eyes following the object to avoid it looking like asleep”(P16). There is an interesting comment from P5, “The robot looks like very bored repeating a boring job.”

The *confused* expression had a good match. 18 out of 20 people (90%) chose correctly, marking it in a mean value of 8.75. Although there were 12 people (60%) choosing it as the *unsure* behaviour with the same mean value of 8.75, the reason might be inferred to be the similar meaning of the words *unsure* and *confused*, depending on the feedback during this and previous segments. Otherwise, there were few scores on the behaviours *thinking*, *surprised* and *sad*, about which people who made the choices thought these behaviours “could be there together with the main behaviour [confused]”, especially when they “looked at the static face for a long time to judge for more descriptions” (P3).

The *sad* expression also got a relatively good match to users’ understanding. 90% participants deemed it a *sad* look accurately, rating it in an average score of 8.56. The reason for the correct match was elaborated, “[the face] looks sad because of the curved eyebrows and eyes looking down” (P9). Meanwhile, there were 8 choices (40%) on the *asleep* behaviour with a noticeable mean value of 7.88, towards which people thought, “the almost closed eyelids make it asleep” (P16). Besides, several scores on the *unsure* behaviour could be interpreted as the robot *unsure* the current situation, which had a similar understanding with the original one in nature.

During the following discussion, all the participants thought the facial expressions were easy for them to make sense. Some novel insights were proposed, for instance, P17 doubted that, “the facial expressions might lead to more focus on the ‘face’ while ignored the working arms”. A consensus was reached among all the participants that adding face to the industrial robot is helpful for the novice user, while followed by one’s suspect that, “it may be helpless to the one who already gets used to the robot” (P17). What’s more, both P5 and P15 advised to add other components, such as texts and sound, to help the operator understand better. Since this experiment focuses more on the understanding of the facial expressions, the

other insights from the discussion part are not discussed in detail.

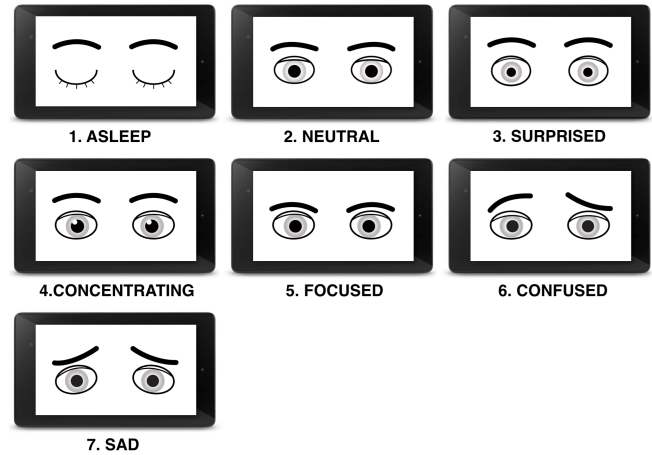


Figure 5: The redesign solution of the set of facial expressions

6. REDESIGN

The next phase followed is adjusting and redesigning the original set of facial expressions based on the results from the first experiment and some theories described above (i.e. Disney’s Twelve Principles of Animation, the wide-eyed theory and the pupillary response theory).

Five facial expressions in the seven were redesigned or made some adjustments, as shown in Figure 5, including *asleep*, *surprised*, *concentrating*, *focused*, and *sad*. The two remaining expressions, *neutral* and *confused*, show a highly accurate match to the behaviours in Experiment I, and have therefore not been redesigned or adjusted.

As for the *asleep* one, there came several solutions tried at an early stage to make the curves more eye-like, such as adding eyelashes in different ways and emphasising the eye sockets. All these solutions were inspired by the Disney’s animation principles and the faces of Disney’s cartoon characters. The second solution was later chosen as the new facial expression to demonstrate *asleep*, which is regarded as more neutral and more like a sleeping face.

In terms of the *surprised* behaviour, the feedback from the first experiment includes that the colour change is tricky, and the “surprised feeling” is not enough. Then the red colour was decided to be removed and the eye sockets were enlarged when showing the *surprised* look, in order to enhance the “surprise feeling”. The redesign of eyes in this facial expression follows the Wide-eyed theory which illustrates that the widened eyes help locate the objects that surprise the robot and reduce the response time for catching operator’s attention quickly.

Concentrating and *focused* behaviours have a similar literal meaning of which people are hard to tell the difference, and their counterpart facial expressions also have the same eyebrows animation that makes the robot look *angry*. In that case, it is necessary to set apart from the two facial expressions based on the context of each. As settled in the original context setting, the *concentrating* behaviour happens when operator teaches the robot, while working robot will show *focused* behaviour. Thus, the solution for the *concentrating* face is animating raised eyebrows and adding white dots of highlight into eyes to show interests and *expecting* to learn new things. The pupillary response theory of human is applied in the *concentrating* face redesign, which presents robot's attention on the operator. Conversely the *focused* face is redesigned as eyebrows dropped down and eyes looking down while following the moving arm instead of closing eyelids, which shows more focus on the working staff and avoid the *asleep* feeling. Accordingly, the two facial expressions are deemed to set more apart from each other while accord to their respective context.

The *sad* face got a good match to the corresponding behaviour, while there is an existent misunderstanding that some people thought the robot looked *asleep*. The reason is analysed to be the closing eyelids. Thereupon, the eyelids are kept opening in the redesigned face to avoid the ambiguous feeling.

7. RESULTS OF EXPERIMENT II

The Experiment II was conducted over two days in April 2016 with 20 participants (named from P21 to P40 below) conforming to the speciality of target users while differ in gender, age and background. All the participants responded positively to the questions asked in the discussion part. The results were analysed in an overall consideration among the quantitative data from the questionnaires, and the qualitative data from participants' comments during tests and insights from the discussion section.

To answer the third research sub-question, a data comparison between Experiment I and Experiment II was conducted and considered as the key phase to inspect whether the redesigned set of facial expressions mapped better to users' understanding than the original set of facial expressions. The statistical analysis began with calculating the number of choices, the mean value (M) and standard deviation (SD) of scores, followed by a Welch's t-test to evaluate the significant difference in the means between the two experiments. Figure 6 shows the data analysis of Experiment II.

Two facial expressions (*neutral* and *confused*) are kept the same in this experiment as with Experiment I, since they are well-matched to the behaviours based on the first experiment data. In Experiment II, both *neutral* and *confused* behaviour got a highly correct map to the



Figure 6: Data analysis of Experiment II.
The number shows the mean value of scores on each facial expression.

behaviours which shows similar results as Experiment I. One people thought, “[the *neutral* face] is like a ‘good morning’ face because it wakes up then starts to expect a new day” (P32). This can be deemed as a positive feedback on the *neutral* behaviour (also on the *asleep* behaviour), since the animation in this segment changes from *asleep* to *neutral*. In addition, data gathered in this experiment further verified that the literature meanings of *confused* and *unsure* are hard to distinguish.

Chart (a) in Figure 7 comparing the survey results of the *asleep* expression before and after redesign, shows most people (95%) correctly choosing the *asleep* behaviour with a prominent result ($M=8.25$, $SD=2.31$), which is better than that before redesign ($M=8.00$, $SD=3.54$). Meanwhile, the choices on the *neutral* behaviour were obviously decreased from 55% to 20%, with a significant change in result ($t(38)=2.20$, $p=0.034$). The results distributed in other options were reduced as well, such as *concentrating*, *confused* and *sad*. Half of the participants mentioned that the robot is *asleep* certainly, and one of them said, “[the robot] is enjoying a tasty cookie” (P30). The results above reveal that the redesigned facial expression (*asleep*) is more understandable for the participants than that before redesign.

As shown in chart (b) of Figure 7, the scores accurately marked on the *surprised* behaviour had an apparent rise from the original expression to the redesigned one ($t(38)=2.17$, $p=0.036$). Additionally the scores distributed in other options were reduced visibly in the scatter diagram, especially the *unsure* behaviour ($t(38)=2.15$, $p=0.038$) and the *confused* behaviour ($t(38)=2.33$, $p=0.025$). However, an increase occurred on the scores of *concentrating* and *focused* behaviours, in terms of which P27 suggested to “make the eyes shift from one person to another” to show surprise caused by the approaching people rather than still focusing on the previous one. It is worth mentioning that the *surprised* facial expression was regarded as an *angry* look because of the red colour in the first experiment, while there was no one mentioning *angry* during the discussion after redesign and only one person (P34) rating this expression as *angry* in a low point (2). It appears that the *surprised* facial expression after redesign maps better to participants’ understanding than the original one. Adding eye-shifting to the face is yet suggested for the sake of locating the concern point.

For the *concentrating* expression, as shown in chart (c) of Figure 7, it is interesting to see that most participants (80%) thought the facial expression meant *angry* and ranked it the first in Experiment I ($M=6.10$, $SD=3.02$), while the choices on *angry* disappeared after redesign and the most conspicuous result suddenly turned to the *surprised* behaviour, with 65% choices and ranking the first ($M=5.05$, $SD=4.02$). Incorrect matches perform in both the two experiments while two different behaviours (*angry* and

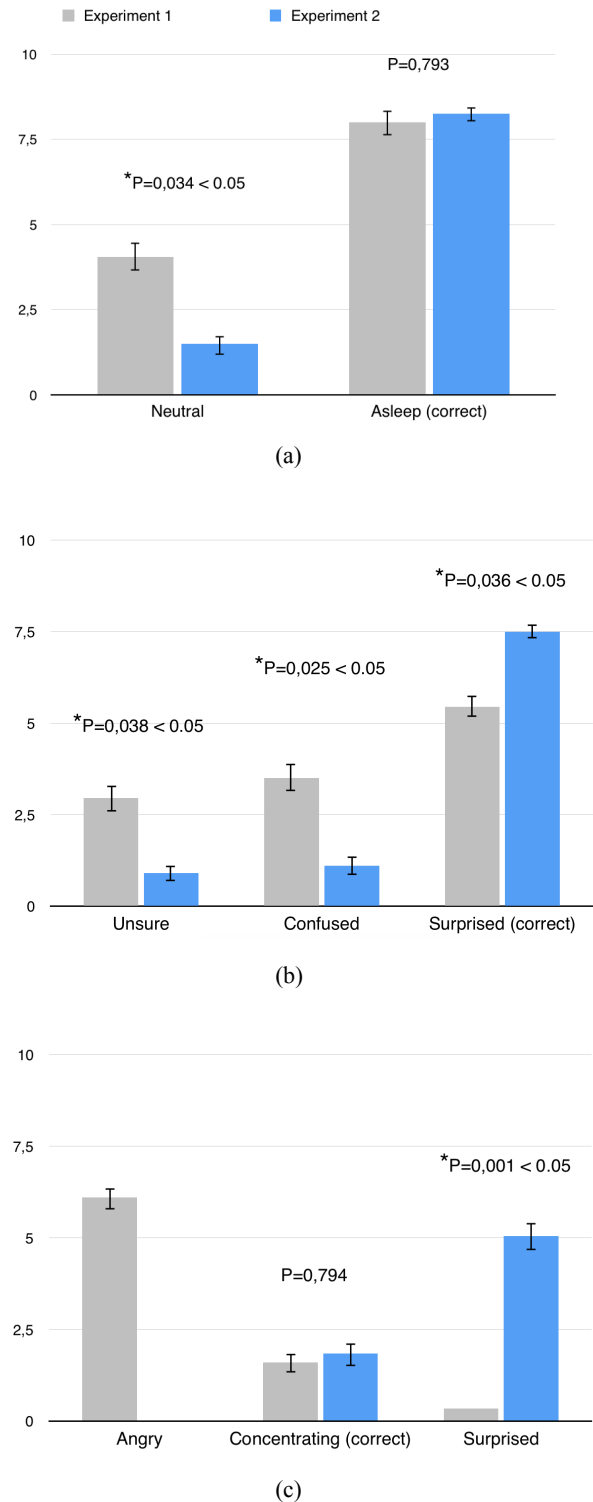


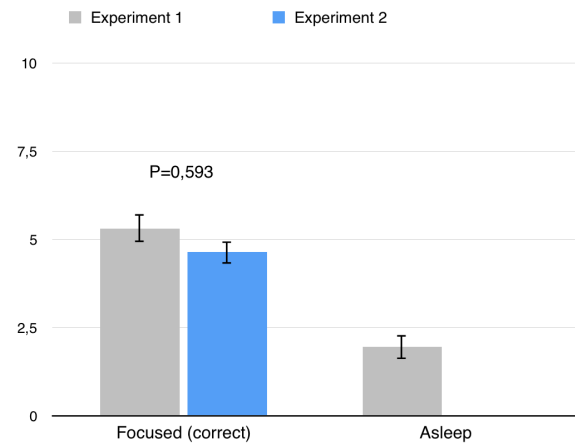
Figure 7: Data comparison in Experiment II. The bar charts show the mean value of the significant data both in Experiment I (grey bar) and Experiment II (blue bar).

surprised) were wrongly matched to the facial expression before and after redesign respectively. The scores marked on the *concentrating* behaviour were invariable between the two experiments. It seems that the redesign of eyebrows animation makes the facial expression less *angry* but more like *surprised* and “interested in something”(P27). As the thing people repeated, a lack of eye-following might be the core reason towards the misunderstandings of this facial expression. If that robot eyes followed the people while learning, the face might be interpreted as a *concentrating* look.

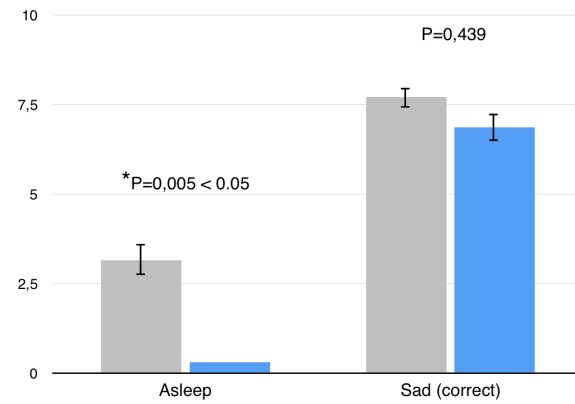
Char (d) of Figure 7 shows the data of the *focused* expression. The scores on the *focused* and *concentrating* behaviours kept the highest after redesign, as well as that in the Experiment I. This persistent result might be due to the ambiguous meaning between the behaviours *focused* and *concentrating*, as participants mentioned when making choices on the questionnaires. Thus, making a clear distinction on the definitions of the two behaviours is necessary. Compared to the scores on the correct match — the *focused* behaviour before redesign (M=5.30, SD=4.16), one more person (70%) chose the *focused* behaviour after redesign with a relatively lower mean value (M=4.65, SD=3.44). The primary reason for this unsatisfactory result seems to be the lack of eyes following the working subject (the same reason of the result in the *concentrating* behaviour). Yet another reason that the facial expression itself doesn't present enough *focused* feelings is considered as well, in terms of which further research on the facial expression presentation is essential in future works. However, the number of choices on the *asleep* behaviour was shown significantly decreased from 5 (25%) before redesign to 0 after redesign, which means the adjustment towards avoiding the sleepy feeling works well.

Based on the data visualised in the chart (e) of Figure 7, correct choices on the *sad* behaviour after redesign (80%) are a little less than that before redesign (90%), while there is no significant difference in the means ($t(38)=0.78$, $p=0.439$). Ruling out of the differences between experiments, it appears that the *sad* feeling of the facial expression keeps well after redesign. Moreover, the scores on the *asleep* behaviour that 40% people found in the original facial expression (M=3.15, SD=4.06) were obviously reduced to zero after resign. Namely, no participant thought this facial expression looked like *asleep* after redesign. Some other oral feedback described the robot as “unsure or confused about what happened” (P25) and “a bit disappointed” (P32). These emotions are appropriate to be seen in the *sad* behaviour, since the behaviour is set in the context of giving up a task or something wrong happened.

During the overall discussion part, a number of perspectives were given out mainly in the 2 respects:



(d)



(e)

Figure 7: Data comparison in Experiment II. The bar charts show the mean value of the significant data both in Experiment I (grey bar) and Experiment II (blue bar).

1) Effectiveness

The comments towards losing the animation were most frequently mentioned, with 30% participants asking to replay the video to find the animations, as P29 said:

“The animation of the facial expression is too tiny to catch my attention and is easy to be missed.” - P29

Besides, the static image of each facial expression seems to be interpreted in multiple meanings by participants, as one commented that,

“I can even match the expression to all these behaviours if I stare at the face for a long time.” - P32

For example, the facial expression that is supposed to express *neutral* was also interpreted as *expecting*, *concentrating* and *thinking* in the two experiments. Thus, it can be inferred that the expression changing process helps people understand the context and make a correct match to the counterpart behaviour. Likewise when the facial expression compared to other facial expressions it can be more understandable (depending on the data comparison). Targeting these problems, some suggestions were proposed by participants such as “make the facial expression more exaggerated and more cartoon-like” (P32), or “add alarm or sound when the face changes to notify the operator” (P27). They also suggested to “add mouth or cheeks” (P23) to the face to make it more expressive, or combine the facial expressions with “some text information” (P27).

2) Value

When it comes to whether or not they think this kind of facial expression is valuable for human-robot collaboration, 95% participants give the positive response. Only one person held a sceptical attitude by considering the uncertainty of future factories.

Another participant emphasised that the value depends on the target and market, because “different targets such as the engineer or the ordinary worker have different needs” (P31), which is the same as different market. For example, this application might be “more user-friendly and valuable in family than that in factories” (P31). What’s more, P27 reminded to take care of annoying people, as for which he advised to provide multiple choices for different users and different usage scenarios.

8. DISCUSSION

To further analyse the results of the laboratory experiments, the sub-questions of this thesis are discussed briefly in this chapter.

8.1 Does the existing set of facial expressions for current industrial robot solutions (i.e. the Baxter robot behaviours) map accurately or not to users’ understanding?

The data and feedback from Experiment I show that three out of seven facial expressions (*neutral*, *confused* and *sad*) map correctly to users’ understanding, while the *sad* expression is concerned with a feeling of *asleep* on account of “the almost closed eyelids” (P16). Another three of them are ambiguous (*asleep*, *surprised* and *focused*), leading to people’s misunderstanding and confusion with other behaviours. The colour usage in the *surprised* expression looks tricky to people. In addition, the *concentrating* expression displays a reversed match to an *angry* face. All the 20 participants thought the facial expressions is easy to understand unanimously, while two of them put forward to provide more information to improve the situation awareness. Applying to Experiment I, it appears that

participants’ understanding towards the current set of facial expressions varies with each individual, but not obviously with the backgrounds and ages of the participants.

8.2 If some of the facial expressions are not understood correctly, how can that be complemented or redesigned to match user’s understanding better?

As described in the Redesign chapter, five of the original facial expressions (i.e. *asleep*, *surprised*, *concentrating*, *focused*, and *sad*) have been redesigned. The redesign process is based on the data analysis and insights towards each facial expression in Experiment I, combining with iterative brainstormings and some theories mentioned earlier, in order to make the facial expressions more expressive and match users’ expectations better. The redesigned facial expressions are being evaluated in Experiment II.

8.3 Does the redesigned version of these facial expressions map to users’ understanding better than the Baxter-like facial expressions?

Based on the data analysis and results comparison in Experiment II, the set of facial expressions shows more understood and effective after redesign. Some of the facial expressions such as *concentrating* and *focused* perform ambiguously either in the original setting or in the redesigned setting, the primary reason of which is the limitation of adding eye-following in this thesis work. The facial expressions are believed to match better to users’ knowledge if eyes direction follows the objects or shifts from person to person.

9. CONCLUSION

After discussing the sub-questions, the conclusion chapter starts with the research question through presenting some principles generated from this thesis research, followed by a consideration of the future work.

9.1 Can people understand industrial robot behaviours better through the redesigned Baxter-like facial expressions?

As discussed in the aforementioned chapters, the affective facial expressions, just as the set of facial expressions in the Baxter robot, are easily understood by people and do add value to the human-robot communication. Moreover, the redesigned version of these facial expressions has proven to more accurately map to users’ understanding. Thus, it is believed that the use of facial expressions for industrial robots is one of the future direction in factories in a 5-year perspective.

In addition, some insights and principles are generated based on this thesis research and can be used in later design

of the facial expressions towards improving the human-robot communication:

1. The cartoon-like robotic interface seems to be more user-friendly and with more “mental safety” than the humanoid robot. Under this precondition, the more human the robot behaves like, the more effective and understandable it can be.
2. Avoiding redundant visual signals such as the improper colour usage that may confuse people is crucial in the future robotic face design. People have different understanding towards some elements, thus an effective human-robot communication should try to identify and handle the discrepancies among areas, cultures, and even individuals.
3. Static images of the facial expressions are often interpreted as multiple meanings. In terms of that enhancing the animations presenting facial expressions change process is important and helps humans aware of the situation and match the behaviour accurately.
4. It is important to make operators notice the facial expression change as the reason mentioned above. The solution for catching people’s attention could be making the facial expressions more expressive and exaggerated, or adding other channels, such as mouth, sound, body movement, etc.

9.2 Future work

The study in this thesis only focused on the tablet facial expressions for industrial robots, which proved to have truly improved the effectiveness of human-robot communication, while it is still restricted and not enough to provide comprehensive information to get users’ prompt responses. By this token, it is clear that the robotic facial expressions studied in this thesis need to be improved to match humans’ expectations better in the future work.

One way for improving the human-robot communication is to keep redesigning the existing facial expressions by following the principles proposed above, to make them more intuitive, and furthermore, fit for the specific robot’s functions and features (this thesis only researched on the Baxter-like facial expressions).

Another direction with enormous potential is combining the facial expression with other channels of interaction or adding multimodal traits, such as adding sound, body gestures and etc. Although the plant environment with many noise may not suit the vocal communication, sound can be still considered as the alternative complementing to the visual signal, in order to show attentiveness in an intimate proximity zone [10]. Otherwise, the body gestures

is believed to reveal more affective state of a robot than the facial expression or even verbal presentation [10], just as the natural face-to-face communication between humans. Hence, we can see that the upcoming researches towards exploring which other traits are useful to facilitate a harmonious human-robot workspace and how to make the components map better to users’ expectation will be a revolution in the future industry.

10. ACKNOWLEDGMENTS

The author of this thesis would like to thank:

- Maria Ralph, supervisor at ABB Corporate Research, for all of her assistance and guidance on experimental design and thesis writing.
- Adam Henriksson, supervisor at ABB Corporate Research, for all of his assistance on design thinking and technical problems.
- Vygandas Vegas Simbelis, supervisor at KTH, for all of his supervision and assistance on the master thesis.
- Mario Romero Vega, examiner at KTH, for all of his guidance on the thesis, especially for his understanding and advice on the confidential problem.

11. REFERENCES

1. Heyer, Clint. "Human-robot interaction and future industrial robotics applications." *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010.
2. Fong, Terrence, Illah Nourbakhsh, and Kerstin Dautenhahn. "A survey of socially interactive robots." *Robotics and autonomous systems* 42.3 (2003): 143-166.
3. Akan, Batu. "Human robot interaction solutions for intuitive industrial robot programming." (2012).
4. Stadler, Susanne, et al. "Anthropomorphism in the factory: a paradigm change?." *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 2013.
5. Rembold, Ulrich, Tim Lüth, and T. Ogasawara. "From autonomous assembly robots to service robots for factories." *Intelligent Robots and Systems' 94. 'Advanced Robotic Systems and the Real World', IROS'94. Proceedings of the IEEE/RSJ/GI International Conference on*. Vol. 3. IEEE, 1994.
6. Fitzgerald, Conor. "Developing baxter." *Technologies for Practical Robot Applications (TePRA), 2013 IEEE International Conference on*. IEEE, 2013.

7. Hoffman, Guy, and Cynthia Breazeal. "Achieving fluency through perceptual-symbol practice in human-robot collaboration." *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*. ACM, 2008.
8. Hoffman, Guy, and Cynthia Breazeal. "Collaboration in human-robot teams." Proc. of the AIAA 1st Intelligent Systems Technical Conference, Chicago, IL, USA. 2004.
9. Hoffman, Guy. "Evaluating fluency in human-robot collaboration." International conference on human-robot interaction (HRI), workshop on human robot collaboration. Vol. 381. 2013.
10. Bethel, Cindy L., and Robin R. Murphy. "Affective expression in appearance constrained robots." *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 2006.
11. Gockley, Rachel, Jodi Forlizzi, and Reid Simmons. "Interactions with a moody robot." *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 2006.
12. Mizanoor, Rahman SM, et al. "Dynamic Emotion-Based Human-Robot Collaborative Assembly in Manufacturing: The Preliminary Concepts." *Workshop on Human-Robot Collaboration for Industrial Manufacturing at RSS'14*. 2014.
13. Jafar, Fairul Azni, et al. "Investigation of human emotional state in human-robot collaboration." *Journal of Computers* 9.3 (2014): 668-677.
14. Nurimbetov, Birzhan, et al. "ChibiFace: A sensor-rich Android tablet-based interface for industrial robotics." 2015 IEEE/SICE International Symposium on System Integration (SII). IEEE, 2015.
15. Reeves, Byron, and Clifford Nass. "How people treat computers, television, and new media like real people and places." CSLI Publications and Cambridge university press, 1996.
16. Thomas, Frank, Ollie Johnston, and Frank. Thomas. *The illusion of life: Disney animation*. New York: Hyperion, 1995.
17. Van Breemen, A. J. N. "Bringing robots to life: Applying principles of animation to robots." *Proceedings of Shapping Human-Robot Interaction workshop held at CHI 2004*. 2004.
18. Ribeiro, Tiago, and Ana Paiva. "The illusion of robotic life: principles and practices of animation for robots." *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 2012.
19. Ellis, C. J. "The pupillary light reflex in normal subjects." *British Journal of Ophthalmology* 65.11 (1981): 754-759.
20. Hess, Eckhard H., and James M. Polt. "Pupil size as related to interest value of visual stimuli." *Science* 132.3423 (1960): 349-350.
21. Lee, Daniel H., Joshua M. Susskind, and Adam K. Anderson. "Social transmission of the sensory benefits of eye widening in fear expressions." *Psychological science* 24.6 (2013): 957-965.
22. Bogue, Robert. "Europe continues to lead the way in the collaborative robot business." *Industrial Robot: An International Journal* 43.1 (2016).
23. Kimura, Hiroshi, Tomoyuki Horiuchi, and Katsushi Ikeuchi. "Task-model based human robot cooperation using vision." *Intelligent Robots and Systems, 1999. IROS'99. Proceedings. 1999 IEEE/RSJ International Conference on*. Vol. 2. IEEE, 1999.
24. Ericsson, K. Anders, and Herbert A. Simon. "Verbal reports as data." *Psychological review* 87.3 (1980): 215.
25. Hoffman, Guy, et al. "Design and evaluation of a peripheral robotic conversation companion." *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015.
26. Duffy, Brian R. "Anthropomorphism and the social robot." *Robotics and autonomous systems* 42.3 (2003): 177-190.
27. Vijai N. Giri. "Nonverbal Communication Theories". *Encyclopedia of Communication Theory*. 2009. <doi: 10.4135/9781412959384.n262>.
28. Albert, William, and Thomas Tullis. *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes, 2013.
29. IFR | International Federation of Robotics. "Provisional definition of Service Robots English." 2012. Web. <<http://www.ifr.org/service-robots/>>. Accessed 9 May. 2016.
30. Dragos, Calin. "3 different humanoid robot head designs to generate facial expressions." *Intorobotics*. 2015. Web. <<http://www.intorobotics.com/3-different-humanoid-robot-head-designs-to-generate-facial-expressions/>>. Accessed 20 Mar. 2016.
31. ABB. "About ABB." 2016. Web. <<http://new.abb.com/about>>. Accessed 27 May. 2016.
32. Jibo blog. "Jibo: Does This Body Make My Head Look Big?" 2015. Web. <<https://blog.jibo.com/2015/08/26/jibo-does-this-body-make-my-head-look-big/>>. Accessed 23 Mar. 2016.

33. RobotsandI. "Have No Fear, the Robot Takeover Here." *Bitlanders*. 2013. Web. <<http://www.bitlanders.com/blogs/have-no-fear-the-robot-takeover-is-here/75976>>. Accessed 12 Feb. 2016.
34. Kismet | MIT. "Kismet, the Robot." 2000. Web. <<http://www.ai.mit.edu/projects/sociable/baby-bits.html>>. Accessed 12 Feb. 2016.
35. Frank Tobe. "Will Be A Pivotal Year for Social Robots." 2016. Web. <<https://www.therobotreport.com/news/2016-will-be-a-big-year-for-social-robots>>. Accessed 2 Feb. 2016.
36. Fawcett, Tom. "An introduction to ROC analysis." *Pattern recognition letters* 27.8 (2006): 861-874.

12. APPENDIX

12.1 Appendix 1: Questionnaire

Here the "Segment #" means the serial number (from 1 to 7) of each segment. In an actual test, the questionnaire includes seven segments, each with a same behaviours list as follows.

The research will remain anonymous.

Thanks for taking your time to improve the human-robot communication.

Your name: _____

Your background: _____

Example

What do you think the robot's facial expression means?

Select and **rate one or more behaviors** you think the robot is presenting.

Behavior



Imperfect match

Perfect match

SEGMENT

Select and **rate one or more behaviors** you think the robot is presenting.

Neutral



Focused



Unsure



Angry



Thinking



Concentrating



Confused



Surprised



Expecting



Asleep



Sad



12.2 Appendix 2: Participants population

	Participant	Gender	Age	Background	Disability	Accuracy	Average of Accuracy
Experiment I	1	Female	41	Control and Analytics	No	57%	61,6%
	2	Female	22	User Experience	No	43%	
	3	Female	46	Software Architect	No	28%	
	4	Female	37	Human Computer Interaction	No	72%	
	5	Male	33	Human Computer Interaction	No	72%	
	6	Female	46	Accounting	No	43%	
	7	Female	44	Embedded System	No	57%	
	8	Female	29	Software	No	72%	
	9	Male	28	Mechanics	No	86%	
	10	Male	27	Interaction Design	No	57%	
	11	Male	29	Mechanical Design	No	43%	
	12	Male	25	Human Computer Interaction	No	57%	
	13	Female	22	Computer Engineering	No	72%	
	14	Female	24	Embedded System	No	72%	
	15	Male	25	Embedded System	No	28%	
	16	Male	26	Automation System	No	57%	
	17	Male	25	Human Computer Interaction	No	86%	
	18	Female	33	Electrical Engineering	No	72%	
	19	Male	26	Mathematics	No	72%	
	20	Male	29	Telecommunications	No	86%	
Experiment II	21	Male	25	Engineering	No	57%	76,1%
	22	Female	26	Electrical Engineering	No	86%	
	23	Male	29	Computer Engineering	No	72%	
	24	Female	23	Wireless Communication	No	86%	
	25	Male	27	Mechatronics	No	72%	
	26	Female	28	Assistant	No	100%	
	27	Male	26	Mechanical Engineering	No	86%	
	28	Male	25	Electrical Engineering	No	72%	
	29	Female	29	Marketing	No	86%	
	30	Male	35	Electrical Engineering	No	43%	
	31	Female	24	Computer Science	No	72%	
	32	Female	32	System Control	No	72%	
	33	Male	25	Engineering	No	86%	
	34	Female	31	Automation System	No	72%	
	35	Female	41	Post-doc Researcher	No	72%	
	36	Male	26	Physics	No	100%	
	37	Male	25	Design Engineer	No	72%	
	38	Female	41	Service Design	No	57%	
	39	Male	48	Intellectual Property	No	86%	
	40	Male	43	Logistics	No	72%	

