Robot Gaze for Communicating Collision Avoidance Intent in Shared Workspaces

Master thesis in the field of study “Computational Engineering” by Li Liu
Date of submission: May 23, 2023

1. Review: Dr.-Ing. Dorothea Koert
2. Review: Dr. Alap Kshirsagar
Darmstadt

TECHNISCHE UNIVERSITÄT DARMSTADT

field of study:
Computational Engineering
Centre for Cognitive Science
IKIDA

Erklärung zur Abschlussarbeit gemäß § 22 Abs. 7 und § 23 Abs. 7 APB der TU Darmstadt

Hiermit versichere ich, Li Liu, die vorliegende Masterarbeit ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Mir ist bekannt, dass im Fall eines Plagiats (§ 38 Abs. 2 APB) ein Täuschungsversuch vorliegt, der dazu führt, dass die Arbeit mit 5,0 bewertet und damit ein Prüfungsversuch verbraucht wird. Abschlussarbeiten dürfen nur einmal wiederholt werden.

Bei der abgegebenen Thesis stimmen die schriftliche und die zur Archivierung eingereichte elektronische Fassung gemäß § 23 Abs. 7 APB überein.

Bei einer Thesis des Fachbereichs Architektur entspricht die eingereichte elektronische Fassung dem vorgestellten Modell und den vorgelegten Plänen.

Darmstadt, 23. Mai 2023

Li Liu

Contents

1 Introduction

1.1 Motivation

In industrial applications, conventional robots have achieved great success because it offers the possibility to combine its high level of automation with human soft skills without further interaction with humans [60]. Collaborative robots (or cobots [13]) have revolutionized the development of robotics by enabling physical human-robot interaction (pHRI) in shared workspaces compared to conventional robots [44]. Relevant studies have shown that collaborative robots are widely used in medical applications, manufacturing, wearable robots, etc. [56].

(a) (b)

Figure 1.1: Figure (a) shows a conventional industrial robot with a cage around it to ensure a safe distance [49]. Figure (b) shows a cobot cooperating with a human user to accomplish a specific task. In this case, there is a potential collision risk [48].

In most conventional robot application scenarios, there is a safe distance between the robot and the human user, as shown in Figure 1.1(a), to ensure collision avoidance [37]. Instead, due to the existence of pHRI, direct physical contact between the robot and the human user is possible, see Figure 1.1(b). For example, learning from demonstrations in the context of reinforcement learning reduces the prior knowledge required by unskilled users to teach robots new skills [55]. This needs to handle possible collisions. On the one hand, the robot needs to adapt its trajectory to the human's movement, and on the other hand, the robot needs to provide the human with sufficient information about its own movement.

One possible solution is to have the robots send signals about their intents that the human user can easily understand and then prepare for their next move. Lemasurier et al. [41] conducted several studies on three light-based intent signals (including gaze, arm light, LED bracelet), three motion-based intent signals (including head pan, forearm movement, gripper movement) and a no-signal (control) condition. The results supported that the average rating of signal noticeability by human participants was significantly higher in both the motion-based and light-based signaling conditions than in the control condition. This means that these signals can help robots communicate their intents and help human participants better perceive the robot's behavior. There are many similar ways to communicate robot intents. Even using projections to indicate the robot's direction of motion has been shown to be effective [57].

However, which signal to use also depends on the particular application scenario. For example, many outdoor environments and production facilities are full of noise that can significantly reduce the effectiveness and reliability of the verbal signal. We notice that many robots use a tablet as their head, which is suitable for displaying gaze signals. With a tablet, the size, shape, direction and other features of the gaze can be flexibly adjusted. The benefits of robot gaze, including improved timing and fluency in handovers and human perception of the robot's navigational intents, have been described in previous studies [19, 18]. To explore how the robot's gaze behaviors should corporate with the robot arm's movement to signal more informative intents, the mechanism of eye-arm coordination should be investigated. Olson et al. [47] found that human users can accurately infer the robot's delivery intents using several well-designed eye-hand coordination patterns inspired by human behaviors.

Obviously, the robot's gaze behavior needs to be further explored and adapted to specific application scenarios. In this thesis, the potential of robot gaze for communicating collision avoidance intent in shared workspaces was investigated. We implemented the temporal scaling method for obstacle avoidance presented by Koert et al. [35] in the context of assembly tasks. This method was developed to improve the safety perception of the human user and the fluency of human-robot interaction (HRI). Furthermore, various gaze behaviors were developed to improve human users' perception of robot intents. Finally, we conducted a user study to find empirical evidence on our topic. Participants performed a collaborative task with the robot and evaluated their preference for the robot's gaze behaviors.

1.2 Structure of this Thesis

This thesis consists of the following chapters:

Chapter 2 provides basic information for understanding the technical aspects of the thesis. In addition, the context of the research problem and a comparison of the current work with previous work are given.

Chapter 3 describes the results of pilot studies and the design process for gaze behaviors.

Chapter 4 presents the design process of the study, including important aspects of empirical research such as hypotheses and detailed experimental methods.

Chapter 5 describes the methods used to analyze the collected data and provides possible interpretations of the results.

Chapter 6 summarizes the findings and discusses future work.

2 Fundamentals and Related Work

This chapter will first introduce the basic terms and methods used in this thesis. Instead of concrete examples, an introduction will be given. Understanding these basics will help readers understand the relevant technical aspects of this thesis. In addition, we will investigate the relevant literature and compare the existing results with an intuitive comparison table.

2.1 Fundamentals

The fundamentals come mainly from the fields of machine learning, statistics, and robotics. Most of them are classical terms or methods that are widely used. The theory of the temporal scaling method and the related concept come from the work of Koert et al. [35]. We recommend readers who are not familiar with these areas to use this section to gain sufficient knowledge to understand this thesis.

2.1.1 Radial Basis Function

A radial basis function (RBF) is a real-valued function ϕ usually used to approximate functions. The value of a RBF depends only on the distance between the input and a fixed center [64]. As a typical representative of RBF, a Gaussian basis function can be represented as

$$\phi(r) = e^{-(\epsilon r)^2} \tag{2.1}$$

where ϵ is used to control the shape of the RBF, $r = \|x - x_i\|$ denotes the Euclidean distance between the input and the center point. As shown in Figure 2.1, a RBF is strictly positive definite.

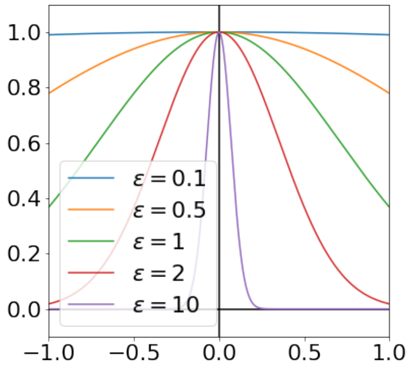

Figure 2.1: Gaussian basis function for several choices of ϵ[64]. As we can see, ϵ controls the shape of the RBF and a RBF is strictly positive definite.

2.1.2 Linear Regression

Linear regression aims to model the relationship between dependent and independent variables. For example, given a data set $\{y_i, x_{i1}, x_{i2}, ..., x_{ip}\}_{i=1}^{n}$, a linear regression model assumes that the relationship between the dependent variable y and independent variables x is linear. Moreover, a noise term ϵ needs to be added to the model. Thus, the linear model can be represented as

$$y_i = \beta_0 + \beta_1 x_{i1} + ... + \beta_p x_{ip} + \epsilon_i, \quad i = 1, ..., n, \tag{2.2}$$

or in matrix notation as

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}, \tag{2.3}$$

where $\boldsymbol{\beta}$ denotes the regression coefficients. Linear regression models are often fitted by minimizing the least squares defined as $RSS(\boldsymbol{\beta}) = (\boldsymbol{y} - \boldsymbol{X\beta})^T(\boldsymbol{y} - \boldsymbol{X\beta})$. Other models minimize a penalized version of the least squares cost function. For instance, the ridge regression (L2-norm penalty) defines the cost function as $(\boldsymbol{y} - \boldsymbol{X\beta})^T(\boldsymbol{y} - \boldsymbol{X\beta}) + \gamma\boldsymbol{\beta}^T\boldsymbol{\beta}$. This can help push the regression coefficients towards zero to improve the performance of the model on new data. Differentiating the cost function by $\boldsymbol{\beta}$ gives the closed-form solution of the linear regression models.

2.1.3 Maximum Likelihood Estimation

Maximum likelihood estimation (MLE) is a widely used machine learning method that estimates the parameters of an assumed probability distribution from observed data. MLE constructs the likelihood function under the assumed statistical model and maximizes it so that the observed data is most likely [61]. The parameter that maximizes the likelihood function is called the maximum likelihood estimate [54].

To illustrate the concept of MLE, suppose there is a random sample data X_1,X_2, ..., X_n. It is assumed that the probability distribution of this sample depends on the parameter θ [45]. This means that the probability density function (PDF) or probability mass function (PMF) of each X_i is $f(x_i|\theta)$. We call the joint PDF or PMF of this sample data $L(\theta)$

$$\begin{aligned} L(\theta) &= P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) \\ &= f(x_1|\theta)f(x_2|\theta)...f(x_n|\theta) \\ &= \prod_{i=1}^{n} f(x_i|\theta) \end{aligned} \tag{2.4}$$

The sample is random, i.e., the X_i are independent. Thus, the second equality holds. MLE finds the parameter θ that maximizes the likelihood function $L(\theta)$.

2.1.4 Bayes' Theorem

Bayes' theorem is a theorem from probability and statistics. With Bayes' theorem, we can represent the probability of an event with prior knowledge of the conditions associated with the event [62]. For example, for a classification task, Bayes' theorem can be described as follows. Given an instance x to be classified. There are K possible classes C_k to which the instance should be assigned. Using Bayes' theorem, the conditional probability can be decomposed as

$$p(C_k|x) = \frac{p(x|C_k)p(C_k)}{p(x)} = \frac{p(x|C_k)p(C_k)}{\sum_j p(x|C_j)p(C_j)} \tag{2.5}$$

2.1.5 Gaussian Mixture Model (GMM)

A Gaussian mixture model is a probabilistic model that fits the data with a mixture of a finite number of Gaussian distributions with unknown parameters. It's general mathematical expression is

$$p(\boldsymbol{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_k) \tag{2.6}$$

where π_k are weights and $\sum_{k=1}^{K} \pi_k = 1$, $\boldsymbol{\mu}_k$ denotes the mean of the k-th Gaussian distribution.

2.1.6 Quadratic Programming (QP)

As a type of nonlinear programming, Quadratic Programming (QP) is the procedure for optimizing mathematical optimization problems with quadratic functions subject to linear constraints on the variables [63]. The standard form of a QP can be formulated as

$$\begin{aligned} \min_{x} \quad & \frac{1}{2} x^T P x + q^T x \\ \text{s.t.} \quad & Gx \leq h \\ & Ax = b \end{aligned} \tag{2.7}$$

where the objective function is convex if and only if P is positive-semidefinite. For convex QP problems, there are several potential solution methods, such as the interior-point method and the active-set method. In this thesis, we used the CVXOPT framework and expected a problem of the standard form shown above, defined by the parameters $\{P$, q, G, h, A, $b\}$.

2.1.7 Probabilistic Movement Primitives

Probabilistic movement primitives (ProMPs) are used to realize a probabilistic interpretation of movement primitives (MPs). Specifically, a ProMP models a distribution over trajectories. As a promising framework for representing and learning MPs, the concept of ProMPs has several advantages, including support for simultaneous activation and the ability to adapt to altered target positions, etc. [50]. In this framework, the joint or Cartesian position of the robot x_t at time step t is given by a linear basis function model

$$x_t = \boldsymbol{\phi}\left(t\right)^T \boldsymbol{w} + \epsilon, \tag{2.8}$$

where $\boldsymbol{\phi}\left(t\right)$ is a n dimensional vector, each element of which is a time-dependent basis equation, $\boldsymbol{w}$ is a weight vector and ϵ is Gaussian noise with zero mean. Ridge regression is used to obtain the weight vector $\boldsymbol{w}$.

Maximum likelihood estimation (MLE) is used to fit a Gaussian distribution over the weight vectors $p\left(\boldsymbol{w}\right) = \mathcal{N}(\boldsymbol{\mu_w}, \boldsymbol{\Sigma_w})$ to capture the variance of the demonstrations. To adjust the execution speed of the movement, a phase variable z is defined as $z = \alpha t$ to decouple the movement from the time signal, where t represents time. With the scaling factor α, the phase z is designed as $z = 0$ when the movement starts and $z = 1$ when the movement ends.

2.1.8 Online Temporal Scaling of ProMPs

It is necessary to adjust the trajectory of the robot to avoid collisions between the robot and potential obstacles. Koert et al. proposed two methods to achieve this goal: online spatial deformation and online temporal scaling [35]. Here, we present the theoretical basis for the online temporal scaling method for adjusting the robot's motion velocity. We define a generalized logistic function as

$$\sigma(\bar{z}) = \delta z_0 + \frac{\delta z_N - \delta z_0}{1 + (1/\varepsilon_{start})exp(m(\bar{z}_c - \bar{z}))}, \tag{2.9}$$

where the scaled phase $\bar{z} = 100z$, δz_0 is the initial phase velocity, δz_N is the resulting end velocity, m is the parameter controlling the velocity change. $\bar{z}_c$ indicates the phase where the phase velocity starts to change, and in this phase the resulting $\sigma(\bar{z}_c)$ will have a slight change due to ε_{start} relative to $\sigma(z_c)$. Thereby, this generalized logistic function encodes smooth velocity adjustment profiles controlled by the predefined parameters. As shown in Figure 2.2, we can generate different acceleration or deceleration profiles by changing the parameters in Equation (2.8).

Using the generalized logistic function, we can calculate the phase velocity δz using

$$\delta z = \delta z_{max} \sigma(\bar{z}), \tag{2.10}$$

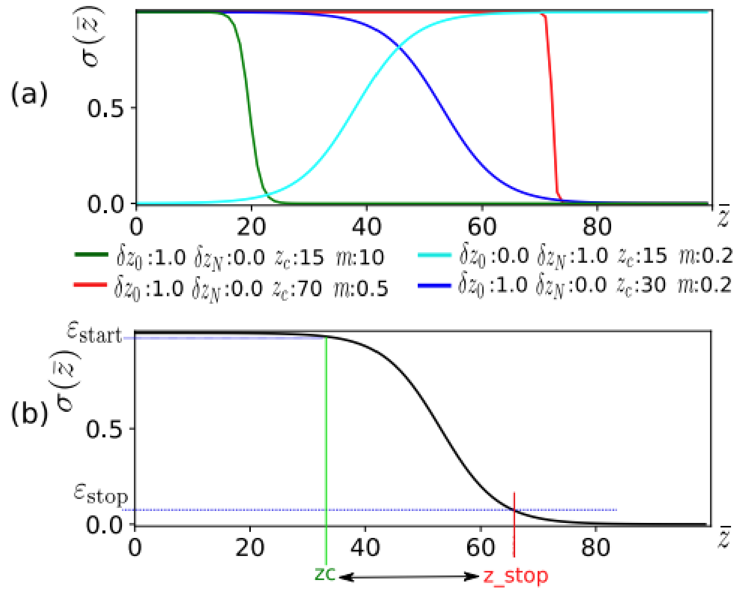

Figure 2.2: (a) By choosing different parameters, we can obtain different acceleration or deceleration profiles. (b) Given a desired stop phase z_{stop} and a optimal deceleration duration we can compute the parameters of the corresponding deceleration profile[35].

where δz_{max} denotes the upper limit for the phase velocity. We define the time at which the phase velocity falls below a predefined small value as $\bar{z}_{stop}$, indicating the stop of the robot motion. If potential collisions with obstacles are detected along the robot trajectory, we can compute $\bar{z}_{stop}$ with an obstacle from a discretized phase vector and then adapt the parameters of the generalized logistic function for a deceleration dependent on the slowing down phase duration

$$\gamma = \bar{z}_{stop} - \bar{z}_c, \tag{2.11}$$

Given $\bar{z}_{stop}$ and the current phase $\bar{z}_n$, the phase duration γ can be determined by solving the constrained optimization problem

$$\begin{aligned} \arg\max_{\gamma} \quad & (\gamma - \gamma_{opt})^2 \\ \text{s.t.} \quad & \bar{z}_{stop} - \gamma > \bar{z}_n \end{aligned} \tag{2.12}$$

where γ_{opt} denotes a desired optimal deceleration duration that should be chosen in advance. According to Equation 2.10, the phase at which the velocity begins to change can be calculated as $\bar{z}_c^* = \bar{z}_{stop} - \gamma^*$, given the optimized γ^*. As described above, we have already gathered knowledge about some of the parameters, including

$$\sigma(\bar{z}_{stop}) = \epsilon_{stop}, \quad \delta z_0 = 1, \quad \delta z_N = 0 \tag{2.13}$$

Thus the optimal slop of the velocity m^* change is obtained by solving Equation (2.8) for m

$$m^* = \log(\frac{\epsilon_{stop}\epsilon_{start}}{1 - \epsilon_{stop}})/(-\gamma^*) \tag{2.14}$$

the resulting values m^* and $\bar{z}_c^*$ are used to update the velocity profile with Equation 2.8. When the potential collision risk disappears, the generalized logistic function will be adapted with an acceleration profile to the original speed.

2.2 Related Work

Compared to conventional industrial robots, collaborative robots (cobots) can achieve collaboration with human users. This eliminates the need for protective cages used to ensure a safe distance, and cobots can be programmed more conveniently with less advanced knowledge than traditional robots [52]. However, a prerequisite for successful collaboration is that safety requirements are met. So far there is no systematic legislation in the field of cobots [52]. Some technical standards refer to the field of cobots, but, they do not include a specification for the design of a safe collaborative shared workspace [26, 27, 25, 28]. There are also some more practical works, e.g., Lasota et al. [40] summarized the commonly used methods for dealing with safety issues due to collisions in the field of HRI, shown in Figure 2.3. In the next chapters, we will focus on safety issues caused by undesired contact between robots and human users in shared workspaces. This thesis relates to three of the four aspects shown in Figure 2.3: control, prediction and psychological consideration.

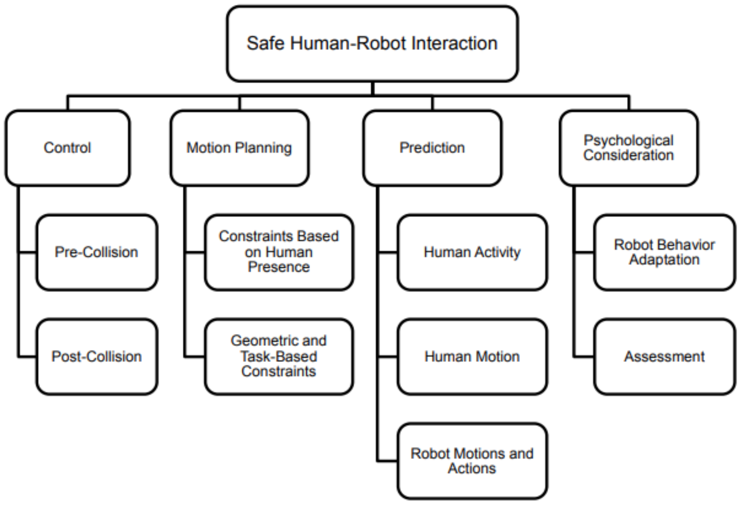

Figure 2.3: The common methods of dealing with safety issues caused by collisions in HRI [40]. This thesis involves three aspects: control, prediction and psychological consideration.

2.2.1 Collision Avoidance for Collaborative Robotics

Collision avoidance and motion planning are two closely related research topics in robotics. The problem to be solved in motion planning is to find a way to guide the robot from the initial state to the target state under the given constraints while avoiding collision between the robot and the environment [24]. Common solutions in this context are usually based on the artificial potential fields (APF) method [16]. Although they are usually easy to implement and have good performance, since they do not involve a generalizable representation of motions, they cannot be generalized for different workspace settings in the scenario of learning from demonstrations. Moreover, methods based on APF are plagued by the global minimum point problem [65]. To obtain generalized motions from human demonstrations, the movement primitives approach modulates the parameters of the robot control policy decomposed from complex motions [50]. Based on movement primitives (MPs), the concept of probabilistic movement primitives (ProMPs) can capture the variance in human demonstrations by working with distributions [50]. Although some literature has extended the ProMPs to apply to interactive environments[34, 12, 43], they are either too computationally intensive or can only deal with static obstacles. Koert et al. [35] developed a goal-based intention prediction model from demonstrations and used the learned model to adapt ProMP trajectories using two approaches: spatial deformation and temporal scaling. Results showed that using intention-aware adaptation, human users can perceive a higher level of safety.

2.2.2 Robot Gaze for Communication

Many ideas in HRI research come from observations of human behavior. A number of studies have revealed that humans infer the intents of their collaborators by observing various behavioral cues [5, 2, 11, 53]. This gives rise to the important concept called joint attention which describes behavior in which attention is shared between people through behavior cues. It is worth mentioning that among the various behavioral cues, the gaze behaviors can intuitively express the intents in the human interaction process in a simple way, which inspires the HRI research.

It is supported by pertinent research that in order for robots to be used in a shared workspace to help humans with their daily tasks, they must be able to engage in joint attention in a manner that is comparable to humans [36]. Inspired by human-human interaction, many research groups have studied the use of robots' active signals in HRI. The findings of Lemasurier et al. [41] confirmed that the noticeability of human user for robot movement can be improved by both the light-based signals (such as gaze, arm light, LED bracelet) and motion-based signals (such as head pan, forearm movement, gripper movement). Furthermore, we also learned that the signal function is directly impacted by the distance between the signal and the human user, and that this effect can have opposing effects on various signal types. It is clear that there are a variety of possible signals that could be used to communicate a robot's intent, but which one is most useful depends on the application scenario. In particular, it is challenging to use verbal signals that humans typically use in noisy factory environments. In this case, light-based or motion-based signals might be more suitable. Some researchers have also studied the use of projections for communicating robot's intent [20]. Although the results indicated that light-based projections can help humans infer the robot's movement direction, a clear drawback of this signal is that it requires human users to move their eyes from the robot to the ground in order to gather information, which could distract human users' attention.

Notably, multiple works have found close connection between gaze and mental focus of attention. One of the important results is the eye-mind hypothesis introduced by Just and Carpenter [31], which provided empirical evidence of the relationship between eye fixation and the information being processed in the mind. The benefits of robot gaze have also been mentioned by many researchers, much like the way humans interact with one another. The use of gaze has been found to be one of the most central and important factors of human interaction that helps to coordinate and confirm the presence of one's interlocutor [32]. Inspired by this, Hart et al. [19] indicated that robot gaze can be used to enhance timing and fluency in handovers by investigating the influence of human-inspired, non-verbal communicative cues during turn-taking tasks. The work of Kshirsagar et al. [38] also provided evidence that well-designed robot gaze can make handover more natural and communicative. According to the research of Boucher et al. [6], human subjects are sensitive to gaze when performing cooperative tasks. More than that, according to the research of Tomasello et al. [59], gaze also aids in indicating the relative position of the robot and human user. The above evidence shows that the robot gaze has the potential to help robots communicate with human users in a shared workspace.

2.2.3 Design of Gaze Behavior to Assist Collision Avoidance

In the preceding subsection, related work has elucidated the importance of gaze behavior as a means of communicating robot intent. However, an appropriate gaze behavior for expressing intent to avoid collision remains an unresolved inquiry. The central aspect here is that the gaze behavior should match the robot's movement to better express the robot's intent, so as to better achieve collision avoidance and improve the subjective interaction experience of the human user.

This problem is related to research into eye-hand coordination (also called reaching ability) of robots. The related work on eye-hand coordination can be divided into two categories: the mathematical approach and the learning approach [10]. The mathematical approach utilizes the robot's forward or inverse kinematics [66] and is well suited to dealing with static environments [30]. On the contrary, the learning approach uses neural networks to build a mapping system from visual perception to hand motor parameters [9].

The problem we are interested in is not the same as most work providing solutions for robot eye-hand coordination. That is, in our context, the robot does not have to have an active machine vision system, which means that the robot's eyes do not perceive the environment. The focus of our study is to use the robot's gaze behavior to communicate robot's intent. There are fewer but similar studies on this topic. Olson et al. [47] designed five gaze behaviors based on observational data obtained from human interactions. Experiments have been conducted in which the robot performs a collaborative task and the information communicated to the humans comes solely from the robots' eye-hand coordination. Depending on the tasks that need to be performed, the robots were assigned specific modes of operation. The authors concluded that human subjects were able to infer the target delivery location with high accuracy. In a study by Neggers et al. [46], different gaze strategies were investigated on human comfort and robot predictability. Although this work is limited to a frontal passing scenario, the results showed that having the robot look at its navigation target is a better strategy than having it face the opposite direction to communicate motion intents to the human user. In addition, results showed that looking in the intended direction is also a suitable strategy for conveying intents.

Even if the results mentioned above are not directly related to collision avoidance, they bring us experiences and insights into the design of the gaze behavior, which we will carry out later.

As a summary of this chapter, Table 2.1 gives an overview of related work, where CA stands for collision avoidance, EA for eye-arm coordination and SW for shared workspace. A checkmark under a topic indicates that the corresponding article mentioned, used, or suggested a theory or method on the topic. On the contrary, a crossmark under a topic indicates that the article does not cover the topic. To the best of the author's knowledge, there is no literature that covers all of these five topics. In fact, our work on designing robot gaze behavior with robotic arm movements to improve the safety of HRI in shared workspaces will bring new insights to the HRI community.

Nr.	Authors & Publication year	Title	Gaze/Intention/ CA/EA/SW
1	Hart et al. 2014.	Gesture, Gaze, Touch, and Hesitation: Timing Cues for Collaborative Work	✓ ✓ × × ✓
2	Lemasurier et al. 2021.	Methods for Expressing Robot Intent for Human–Robot Collaboration in Shared Workspaces	✓ ✓ × × ✓
3	Hart et al. 2020.	Using Human-Inspired Signals to Disambiguate Navigational Intentions	✓ ✓ ✓ × ×
4	Mwangi et al. 2018.	Dyadic Gaze Patterns During Child-Robot Collaborative Gameplay in a Tutoring Interaction	✓ ✓ × × ✓
6	Lukic et al. 2012.	Learning Coupled Dynamical Systems from human demonstration for robotic eye-arm-hand coordination	× ✓ ✓ ✓ ×
7	Lukic et al. 2014.	Learning robotic eye–arm–hand coordination from human demonstration: a coupled dynamical systems approach	× ✓ ✓ ✓ ×
8	Duarte et al. 2018.	Action Alignment from Gaze Cues in Human-Human and Human-Robot Interaction	✓ ✓ × × ✓
9	Olson et al. 2020.	Human-Inspired Robotic Eye-Hand Coordination Enables New Communication Channels Between Humans and Robots	✓ ✓ × ✓ ×
10	Chao et al. 2018.	Enhanced Robotic Hand–Eye Coordination Inspired From Human-Like Behavioral Patterns	× × × ✓ ×
11	Admoni et al. 2017.	Social Eye Gaze in Human-Robot Interaction: A Review	✓ ✓ × × ✓
12	Boucher et al. 2012.	I reach faster when I see you look: gaze effects in human–human and human–robot face-to-face cooperation	✓ ✓ × × ✓
13	Liu et al. 2022.	A Control Strategy of Robot Eye-Head Coordinated Gaze Behavior Achieved for Minimized Neural Transmission Noise	✓ ✓ × × ×
14	Wang et al. 2019.	Gaze awareness improves collaboration efficiency in a collaborative assembly task	✓ ✓ × × ✓
15	Raković et al. 2022.	The Gaze Dialogue Model: Nonverbal Communication in HHI and HRI	✓ ✓ × × ✓
16	Duarte et al. 2018.	Action Anticipation: Reading the Intentions of Humans and Robots	✓ ✓ × × ✓

Continued on next page

Continued from previous page

Nr.	Authors & Publication year	Title	Gaze/Intention/ CA/EA/SW
17	Boschetti et al. 2022.	The influence of collision avoidance strategies on human-robot collaborative systems	× ✓ ✓ × ✓
18	Gasparetto et al. 2015.	Path planning and trajectory planning algorithms: A general overview	× × ✓ × ✓
19	Hwang et al. 1992.	Gross motion planning—a survey	× × ✓ × ✓
20	Zhang et al. 2010.	Dynamic artificial potential field based multi-robot formation control	× × ✓ × ✓
21	Hata et al. 2012.	Target object announcement combining robot gaze and augmented hand	✓ ✓ × ✓ ✓
22	Kshirsagar et al. 2020.	Robot gaze behaviors in human-to-robot handovers	✓ ✓ × × ✓
23	Neggers et al. 2022.	Effect of Robot Gazing Behavior on Human Comfort and Robot Predictability in Navigation	✓ ✓ × × ×
24	Moon et al. 2014.	Meet Me where I'm Gazing: How Shared Attention Gaze Affects Human-Robot Handover Timing	✓ ✓ × × ✓
25	Chao et al. 2012.	A developmental constraint driven approach to developmental robotic hand-eye coordination	× × × ✓ ×
26	Zhou et al. 2016.	Learning Visuomotor Transformations and End Effector Appearance by Local Visual Consistency	× × × ✓ ×

Table 2.1: Overview of related work. CA denotes collision avoidance, EA denotes Eye-arm coordination and SW denotes shared workspace. A checkmark under a topic indicates that the corresponding article mentioned the topic or used/proposed a theory or method related to the topic. On the contrary, a crossmark under a topic indicates that the article does not cover the topic. To the best of the author' knowledge, there is no literature that covers all of these five topics.

3 Gaze Behaviors for Communicating Collision Avoidance Intent

The gaze behaviors investigated in this study were mainly based on the observation of human-human interactions and related literature. However, there are many possible combinations of human-like behaviors that can be used as the robot gaze. In addition, not all gaze behaviors are suitable for our application. Therefore, we conducted two pilot studies and collected feedback from participants to select a set of behaviors that make sense to human cognition as an iterative improvement method.

3.1 Problem Statement

In this thesis, we need to find a suitable robot gaze behavior to communicate robot intent to avoid potential collision. The concrete scenario is that a robot and a human user cooperate to complete a series of assembly tasks. The robot is fixed on the base, so our concern is the collision between the robot arm and human hand. According to the observation of human-human interaction, the robot's gaze behavior plays an important role in this context. We need to design several potentially representative gaze behaviors at first, then compare these behaviors, analyze the differences in their effects on the execution of collaborative tasks and the behavior of human users.

The initial selection of gaze behaviors was mainly based on the feedback from participants and responses to experimenter's questions. The questions were: Did you notice the robot's gaze behavior? How many gaze behaviors did you notice? Do you think it necessary to change the existing experimental setup to improve your overall experience when interacting with the robot?

3.2 Experimental Setup

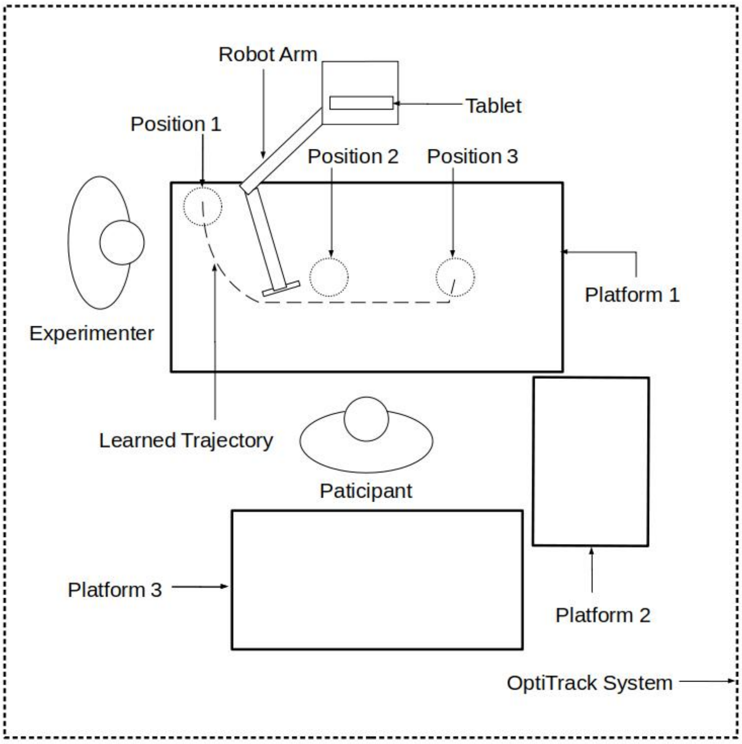

Figure 3.1: The experimental facilities mainly included a Franka Emika Panda robot arm with an end effector, an OptiTrack motion tracking system with 12 camera, a pan-tilt unit (PTU) with a SAMSUNG S6 tablet, and three tables. This figure only shows the forward trajectory of the robot arm, the backward trajectory is similar in shape to the forward trajectory.

As shown in Figure 3.1, the experimental facilities consisted of a Franka Emika Panda robot arm with an end effector, a pan-tilt unit (PTU) with a SAMSUNG S6 tablet, and an OptiTrack motion tracking system with 12 cameras to track the positions of the end effector, the human hand and the human head. Boucher et al. have shown us that using the combined movement of the eyes and the head to express intent is a good way [6], so we selected a static image of the eyes displayed by the tablet as the robot's face and combined it with the mechanical movement of the PTU. Participants used a glove and a hood to attach the marker of the OptiTrack system to their hand and head. The robot arm carried objects from position 1 to position 3 on platform 1 along a Cartesian trajectory learned from three demonstrations. The reciprocating movement of the robot arm consisted of two parts: the forward movement to deliver objects from position 1 to position 3 and the backward movement to return to position 1. The two parts were each learned from three demonstrations. The human user assembled the objects collected from position 2 and position 3 on platform 2 and then placed the assembled objects on platform 3. Position 1 is fixed in advance, the experimenter stood next to position 1 and position 2 to supplement objects.

The trajectory can be learned from joint space or Cartesian space and conditioned by via points using ProMP introduced in Subsection 2.1.7. In our experiments we used the learned mean trajectory from Cartesian space. See Figure 3.2. In addition, if the trajectory of the robot arm was occupied by an obstacle, the robot arm would use the temporal scaling method described in Subsection 2.1.8 to adjust the velocity.

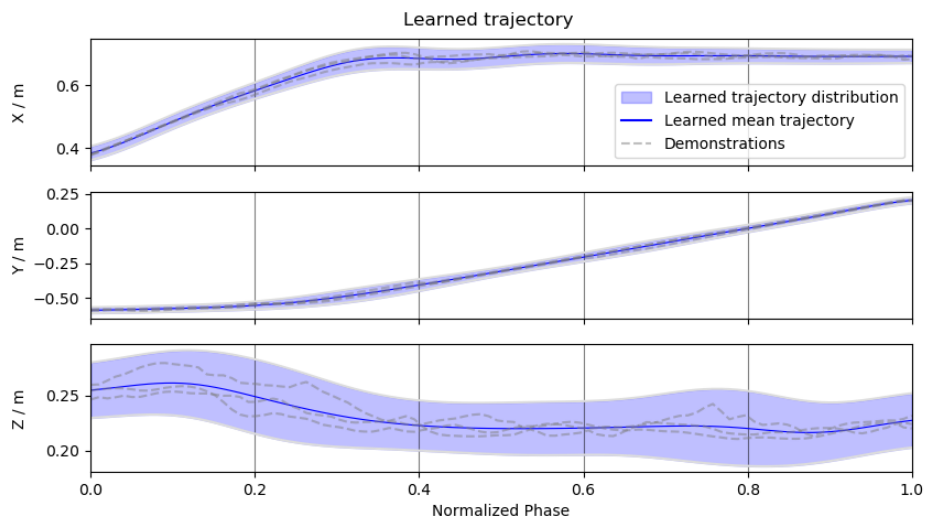

Figure 3.2: The trajectory was learned from Cartesian space using ProMP. The light blue areas represent the learned trajectory distribution. The gray lines represent the 3 demonstrations. As the trajectory of the robot arm, we used the mean value represented by the blue lines.

3.3 Gaze behavior first version

In the first pilot design, the pan-tilt unit (PTU) and the tablet were mounted on top of the robot as shown in the left half of Figure 3.4 because we thought it made the robot look more anthropomorphic. The participant was asked to stand up when interacting with the robot in order to avoid the robot arm interfering with their view. Based on observations of human behavior and inspired by the study of Neggers et al. [46], we thought the robot's gaze toward the motion destination was suitable in safe condition. But when collision risk was present, the robot should exhibit different gaze behavior to communicate with human user. For this purpose, we chose three gaze behaviors: gazing toward the human head, nodding in the direction of the human head and shaking head in the direction of the human head. The state of the robot arm and a distance-based metric, shown in Figure 3.3, determined which behavior the robot will exhibit.

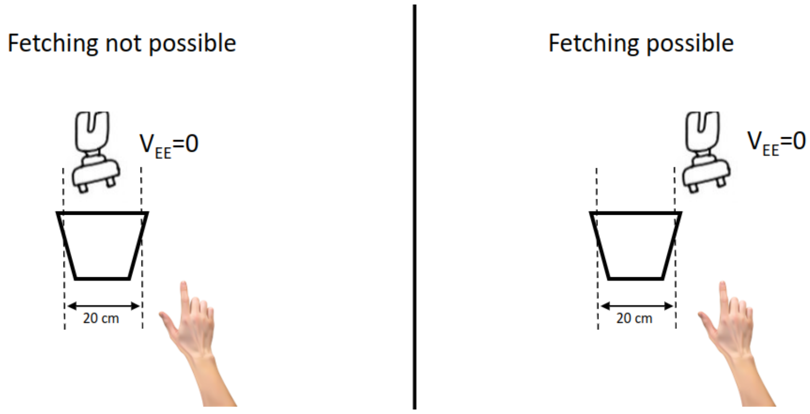

Figure 3.3: Distance-based metric used to determine whether it is possible for the human user to grab the object without colliding with the robot arm after the robot arm has already stopped. In our design, if the end effector of the robot arm is within ten centimeters of the object's pickup location, it's difficult to grasp the target without collision.

Based on the above considerations, we designed the following gaze behaviors:

Condition A: The robot gazes toward its motion destinations.

Condition B: The robot gazes toward its motion destinations. After the robot arm stops because of an obstacle on the trajectory, the robot will gaze toward the human head.

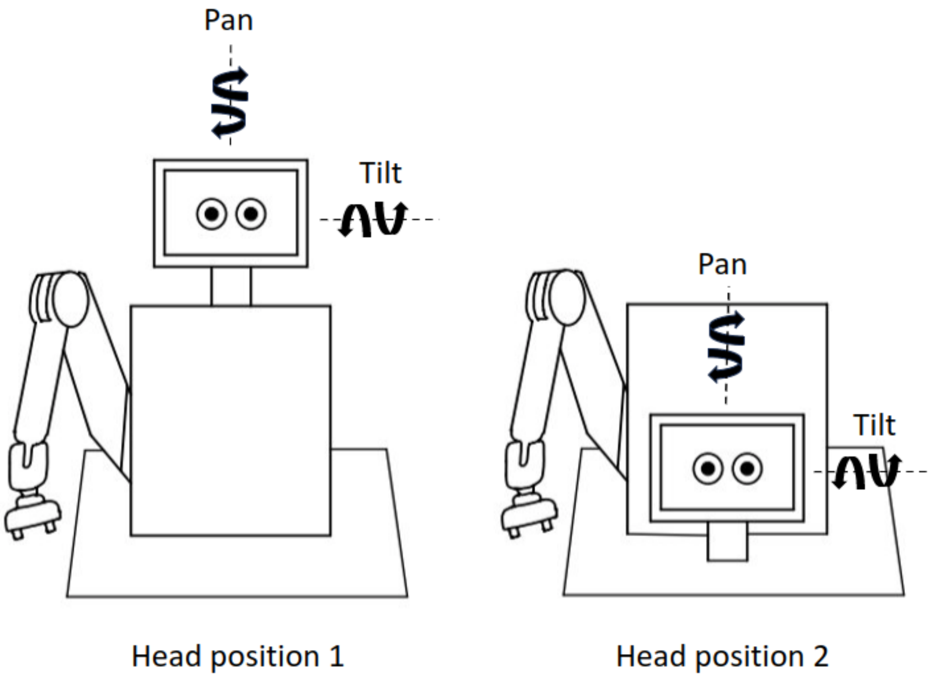

Figure 3.4: Our study included two positions of robot head. When the robot head was on the top (head position 1), the robot looked more anthropomorphic, while the gaze behavior seemed to be more noticeable when the robot head was on the table (head position 2), according to the participant's feedback.

Condition C: The robot gazes toward its motion destinations. After the robot arm stops because of an obstacle on the trajectory, the robot will nod toward the human head.

Condition D: The robot gazes toward its motion destinations. After the robot arm stops because of an obstacle on the trajectory, the robot will nod toward the human head if fetching objects is still possible (shown in the right half of Figure 3.3), or shake the head toward the human head if fetching objects is no more possible (shown in the left half of Figure 3.3).

3.4 Gaze behavior second version

The participant of the first pilot study pointed out a shortcoming in the first design. In this design, the robot only started performing the gaze behavior after the robot arm had stopped. In addition, because the participant was standing at the same height as the robot, the robot must first look up at first before shaking its head or nodding. Added to this is the reaction time of the mechanical components of the pan-tilt unit (PTU). When the participant looked at the robot's head, the robot often had not yet started the gaze behavior. Due to the existence of the assembly task, the participant usually did not wait for the robot to complete its behavior, let alone try to understand it, but continued with the assembly task.

To work around this issue, we added a human hand detection function (pre-emptive collision gaze) to the existing design by defining two risk region around two object's pick-up location, as shown in Figure 3.5. If the human hand was detected, the robot would look up even if the robot arm was still moving. When the robot arm was stopped, the robot head was ready to nod or shake. This made the robot react faster. In addition, we thought it would be interesting to adjust the experimental conditions to compare the difference between head shaking and head nodding.

Based on the discussion above, we adjusted the gaze behavior as follows and conducted a second pilot study:

Condition A: The robot gazes toward its motion destinations.

Condition B: The robot gazes toward its motion destinations. When collision risk exists, the robot will look up. After the robot arm stops because of obstacle on the trajectory, the robot will gaze toward the human head if fetching objects still possible, or shake head toward the human if fetching objects not possible.

Condition C: The robot gazes toward its motion destinations. When collision risk exists, the robot will look up. After the robot arm stops because of obstacle on the trajectory, the robot will nod toward the human head if fetching objects still possible, or gaze toward the human if fetching objects not possible.

Condition D: The robot gazes toward its motion destinations. When collision risk exists, the robot will look up. After the robot arm stops because of obstacle on the trajectory, the robot will nod toward the human head if fetching objects still possible, or shake head toward the human if fetching objects not possible.

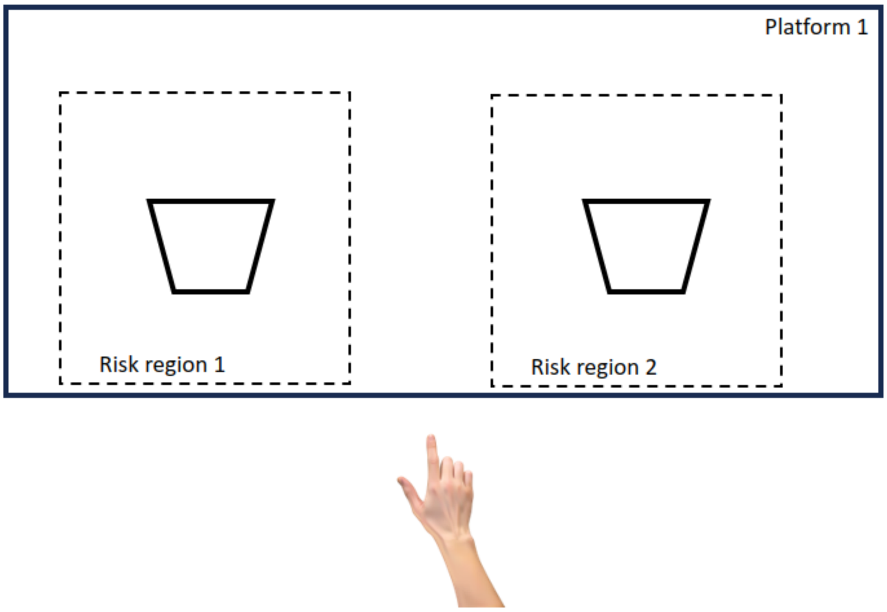

Figure 3.5: We defined two square risk regions with sides of 40 cm. If the human hand enters one of these two regions and the robot arm is also moving towards this region, we believe there is a risk of collision at this time.

3.5 Gaze behavior third version

The participant of the second pilot study provided valuable feedback on the new design. First, since the participant performed the collaborative work while standing, he always had to look down to grab objects. This made it difficult for the participant to pay attention to the robot's gaze behaviors. As a workaround, we asked the participant to sit down on a swivel chair. Second, with this design, in most cases the robot would gaze toward the two destinations on the left or right. When there was a risk of collision, it still took some time to change direction from either side toward the human head (around the middle of the robot head's pan angle), which still caused the participant to feel some delay. To solve this problem, we had the robot look forward when there was no risk of collision. In addition, the participant commented that the head nodding behavior was less noticeable than head shaking behavior. So we reduced the proportion of head nodding behavior in the design and left it in condition D only.

The design of gaze behavior before the start of the experiment was as follows:

Condition A: The robot always looks forward.

Condition B: The robot looks at the human when collision risk exists, otherwise look forward.

Condition C: The robot gazes forward. When collision risk exists, looks at the human. After the robot arm stops because of an obstacle on the trajectory, shakes its head towards the human.

Condition D: The robot nodes toward the human head. When collision risk exists, looks at the human. After the robot arm stops because of an obstacle on the trajectory, the robot nodes toward the human head if fetching objects is possible, or shakes its head toward the human if fetching objects is not possible.

4 Hypothesis-Driven Experiments

Unlike social science research and robotics research, HRI research involves at least two interacting components: the human and the robot. While the participants in HRI appear to be relatively well-defined, the research scope of HRI is broad and interdisciplinary. Some of the work, referred to as robot-centric work, focuses on the technical aspects of the robot with the goal of improving the functionality of the robot itself for interacting with humans. While user-centered work studies the influence of humans on HRI outcomes, the focus is on the way humans perceive and interact with robots in different contexts, as well as empirical studies in traditional social sciences [3].

Empirical studies have become standard across the broad spectrum of HRI research. Rigorous empirical studies are essential to obtain valid conclusions about the performance of new methods [22] and often contain a number of important sessions. For example, the researchers need to specify the context in which HRI takes place, as this helps clarify the application scenarios and the importance of the research results. In addition, the context in which human-robot interactions take place is also an indispensable topic that strongly influences the results.

There are some research methods that are recommended as a guideline for rigorous experimental design in the relevant literature. These methods have some sessions in common. For example, the experimenter needs to find a research question that prevents the author from getting lost in the many details of the experimental design at first. Additionally, statistical tests are typically an essential session that can help researchers verify the accuracy of test results. As shown in Figure 4.1, Gravetter and Forzano proposed a closed-loop research method that uses the result to modify, refine, or extend the original research idea [17]. Another method proposed by Hoffmann et al. [22] does not have this feedback mechanism, but its overall concept is consistent with the previous one, as shown in Figure 4.2. In fact, we went through a similar feedback process with the help of the pilot studies. In this chapter we will describe the empirical study design that follows the main ideas derived from the two methods.

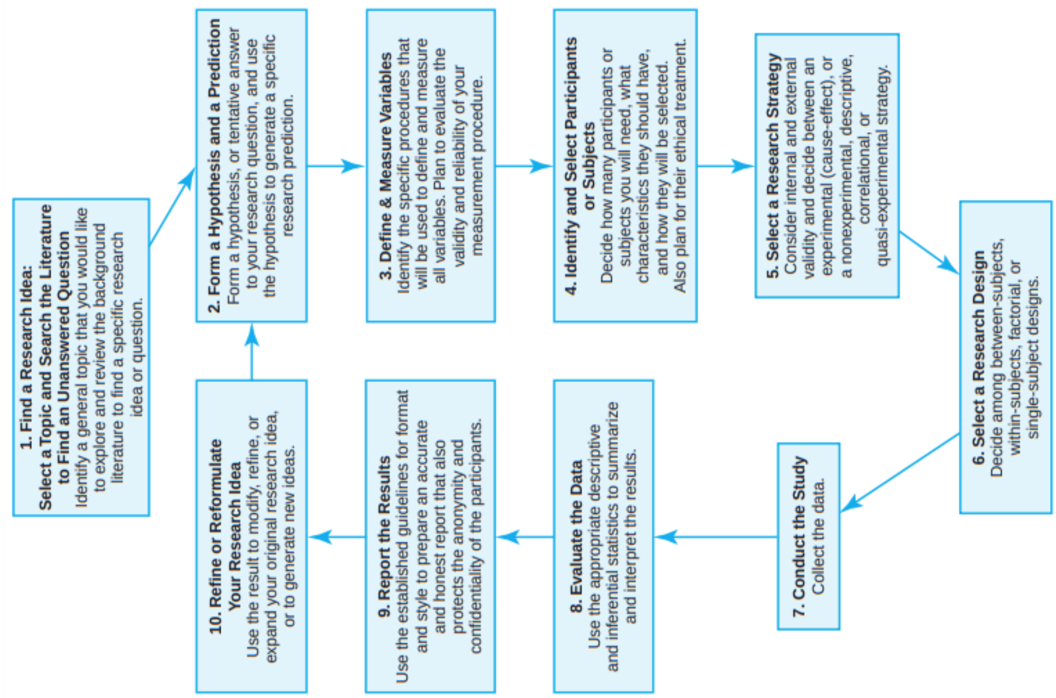

Figure 4.1: The method recommended by Gravetter and Forzano [17] uses a closed-loop mechanism to modify and refine the original research idea.

4.1 Research Questions

In this thesis, we are particularly interested in the relationship between the design of robot gaze behavior and human perception of the robot's intent. Similar to humans performing collaborative tasks, we can expect that if the robot can communicate the motion intent well, the collaborative task will be performed more smoothly, and the collision between human and robot will also be reduced. Therefore, we begin our empirical study with two research questions:

1. To what extent, if any, will well-defined robot gaze behavior lead human users to anticipate the robot's intent?

2. Will human users get any benefit with a robot equipped with well-defined gaze behavior?

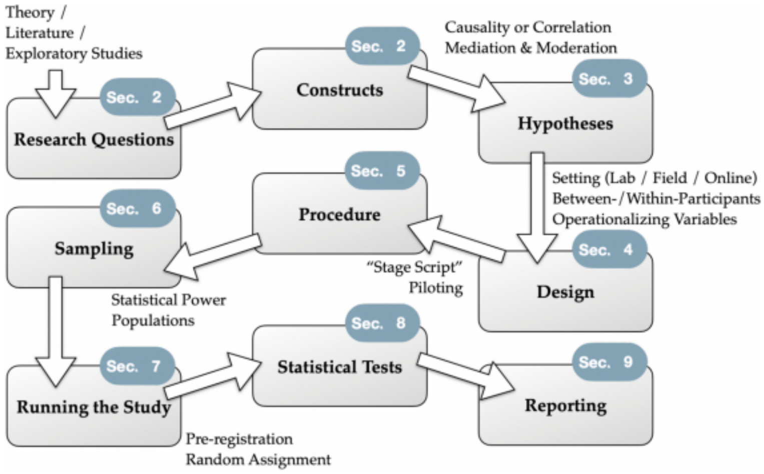

Figure 4.2: The method proposed by Hoffmann et al. [22] does not have a feedback mechanism like the method recommended by Gravetter and Forzano [17]. However, the overall concepts of the two methods are the same.

4.2 Constructs and Hypotheses

Constructs are representations of the concepts of interest in a study. In empirical studies, researchers are concerned with finding the relationships between constructs. In our context, the constructs include: gaze behavior, human's anticipation of danger, intent communication of robot, human-perceived competence of the robot, human discomfort and efficiency of the collaborative task. The gaze behavior contains the four conditions proposed in Section 3.5.

Hypotheses are affirmative statements of relationships between constructs that we can either support or refute. Based on the information obtained from the relevant literature and the observation of human-human interaction, we propose the following hypotheses:

Hypothesis 1: Human users better anticipate danger working with the robot in condition D than in other conditions.

Hypothesis 2: Robot's gaze behavior in condition D better communicates the robot's intent to the human user than gaze behaviors in other conditions.

Hypothesis 3: Gaze behaviors in conditions B, C, and D improve the human-perceived competence of the robot compared with gaze behavior in condition A (control group).

Hypothesis 4: Gaze behaviors in conditions B, C, and D reduce the discomfort of human users compared with gaze behavior in condition A (control group).

Hypothesis 5: Gaze behavior in conditions D reduce task execution time.

4.3 Design of The Study

After the definition of research question and constructs, we have proposed five hypotheses. Now we start designing the study.

4.3.1 Study Context

As shown in Figure 4.3, there are three ways to test the hypotheses: testing in a laboratory, testing in the field, or testing on the Internet. Testing in a laboratory avoids the problem of controlling for confounding variables when testing in the field. Although the characteristics of field studies conducted in daily environments create conditions for research external and ecological validity, in our research scenario, the interaction between the robot and the human will take place in a specific indoor workspace, and the base of the robot will be fixed, what makes fields studies difficult to implement. The third option, testing on the Internet, has become more and more popular in recent years. A common practice in HRI is to show participants videos of humans interacting with robots and questionnaire about their subjective feelings. With this option we can collect data faster while the external validity is lower and we lose control over online participants. In our study, we will mainly use the option of testing in a laboratory.

4.3.2 Between- and Within-Participants Designs

In a between-participants study, each participant is randomly assigned to a group to experience a variation in experimental conditions. We need to compare these different groups of participants to find out the relationships between the constructs. In contrast, in a within-participants design, each person experiences more than one experimental

(a) (b) (c)

Figure 4.3: Figure (a) shows an example of testing in a laboratory [19], a human is collaborating with a robot to complete a puzzle. Figure (b) shows a hallway in which a field study will take place [18]. As an example of testing on the internet, Figure (c) shows a human conducting an online survey [8].

condition. What we need to compare are the different experiences of each participant. An image explanation is shown in Figure 4.4.

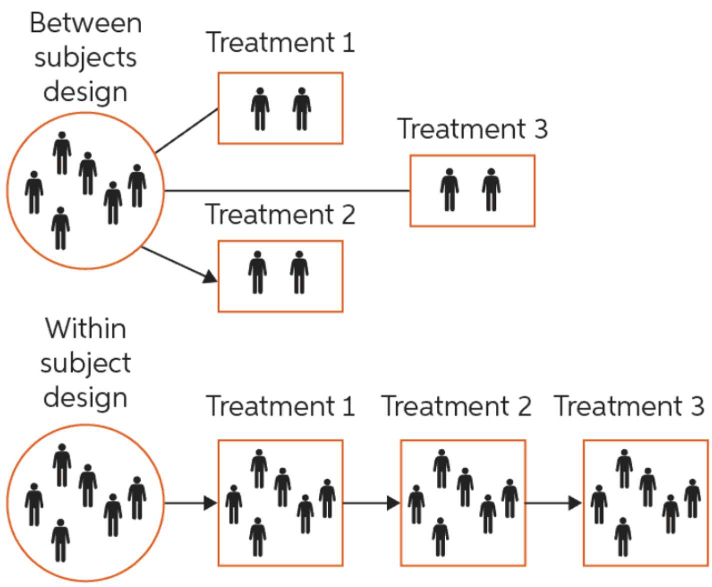

Figure 4.4: The difference between between-participants design and within-participants design is that in between-participants design, participants will only experience one experimental condition.

Although within-participants design suffers from the order effect and the novelty effect, in our scenario we want to take advantage of its low requirement for the number of experiment participants. As a workaround to mitigate the two negative effects, we can use a method called counterbalancing to randomize the order of the conditions that the participants experience.

4.3.3 Operationalizing Constructs into Variables and Measures

Operationalization means converting constructs into specific things that we can manipulate and measure. However, this is not an automatic process as a construct can be operationalized in many different ways. In our study, we can operationalize the efficiency of collaborative task as task execution time, which means the total time spent by participants in completing the collaborative task. For other constructs there aren't very intuitive measures, for which we need to design questionnaires to measure the response of the participants. Bartneck et al. proposed a standardized measurement tool called the Godspeed

Scale which is widely used in HRI with five consistent questionnaires using five-point Likert-type scales [4]. However, the five key concepts of anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety are not compatible with our study. Based on the Godspeed Scale, Carpinella et al. developed the Robotic Social Attribute Scale (RoSAS) [7], which contains exactly the concepts and items that interest us. We can use these items to measure human's judgments of the human-perceived competence of the robot and discomfort. The remaining two constructs of anticipation of danger and intent communication can be measured using two subjective item that we designed ourselves. The four questionnaires used in our study are shown in Table 4.1.

Questionnaire 1: ANTICIPATION OF DANGER										
"I was able to anticipate potential collisions with the robot"										
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
Questionnaire 2: INTENT COMMUNICATION										
"The robot communicated its intent clearly"										
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
Questionnaire 3: PERCEIVED COMPETENCE										
Using the scales provided, how closely are the words associated with your impression of the robot?										
	Knowledgeable									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Interactive									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Responsive									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Capable									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Competent									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Reliable									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
Questionnaire 4: DISCOMFORT										
Using the scales provided, how closely are the words associated with your impression of the robot?										

	Aggressive									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Awful									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Scary									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Awkward									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Dangerous									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree
	Strange									
Strongly disagree	1	2	3	4	5	6	7	8	9	Strongly agree

Table 4.1: We used four questionnaires in our study to measure human's feelings. Questionnaire 1 and questionnaire 2 were designed by us. Questionnaire 3 and questionnaire 4 originate from the work of Carpinella et al. [7].

4.4 Study Procedure

After entering the experimental laboratory, the participants read the informed consent form at first and then signed it if they agreed to it. They then answered general questions about their gender, age range, and familiarity with collaborative robots. Participants then generated an identification code to track their data using a predefined rule. This ensured the anonymity of the participants and at the same time enabled the participants to request the deletion of personal data in special cases. The randomization process was then carried out. Participants were assigned an order of the four experiments to be performed using an integer generated from a website [51]. Twelve participants drew twelve sequences from a total of twenty-four experimental sequences without repetition.

The experimenter then verbally described the task and experiment procedure to the participants. Afterwards, participants wore the glove and hood of the OptiTrack system described in Section 3.2 and completed an approximately three-minute practice session with the robot without robot gaze behavior to become familiar with the learned trajectory, assembly task, and collision detection behavior of the robot. When real study began, participants were asked to complete 30 assembly tasks in each experiment. After each experimental condition, the participants first expressed three feelings about the robot's behavior and then filled out the questionnaires described in Subsection 4.3.3. After the final experiment, the participants also answered five final questions about the four experimental conditions and provided additional feedback. The experimenter used a computer program to record the position information of the end effector of the panda arms, the human hand and the human head during the experiments. In addition, the experiments were recorded by a camera.

Each experiment lasted about 8 minutes, depending on the pace of each participant. The total time to complete the study was approximately 60 minutes. After completing all experiments, the experimenter checked whether the questionnaire was filled out completely and whether the data was saved correctly. Finally, the experimenter revealed the true purpose of the study to the participants and asked the participants if they would like to withdraw their consent. If so, the collected data will be deleted. If not, the experimenter would escort the participant out of the lab. Figure 4.5 describes the study procedure.

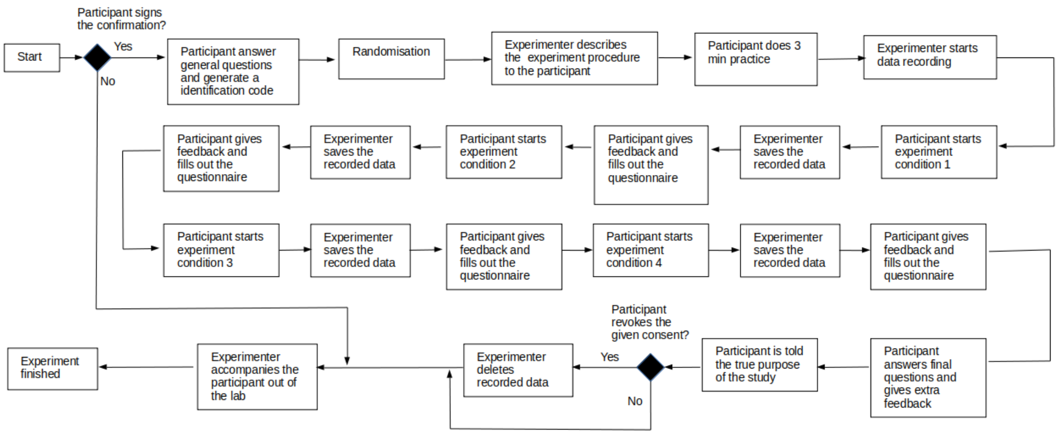

Figure 4.5: The study procedure. A participant may refuse to sign the consent form or withdraw consent after knowing the true purpose of the study. In both cases, no data will be collected for this experiment.

5 Results and Discussion

In this chapter, we will analyse the collected data using ANOVA and correlation analysis with the JASP software [29]. After four completed experiments, the participants reported that human's perception of robot gaze behavior could be improved if the robot head was placed on the table. So we changed the position of the robot head as shown in the right half of Figure 3.4. Given that there are two independent variables (position of the robot head, gaze conditions), the two-way repeated measures ANOVA with a significance level of .05 was used. With a total of 25 participants, we could detect an effect size of d=0.50 with a power of .80 at a significance level of .05 (calculated with the G*Power software [14, 15].) However, due to time constraints and the purpose of the thesis as a pilot study, 12 participants finally took part in the experiment.

In our study, we performed Mauchly's sphericity test as sphericity test and Levene's test as homogeneity test, with a significance level of .05. It should also be noted that the data collected is unbalanced (4 samples for the robot head on top of the robot and 8 samples for the robot head on table). Unequal sample sizes can cause problems, including unequal variances between samples and loss of power. According to the description of Keppel and Wickens [33], there is no good rule of thumb to determine exactly when these problems occur.

Performing an ANOVA on unbalanced data involves the selection of the approach for computing sums of squares (SS). Traditionally, there are three commonly used approaches in this context called Type I-III using the designations from SAS (Statistical Analysis System). According to the conclusions of a number of studies [58, 21], Type III analysis is usually preferred and is the default method for calculating SS by major statistical software, while Type I analysis is strongly influenced by the order of the factors and Type II analysis is based on the assumption that the interactions between the factors are negligible. However, Lewsey et al. [42] claimed that Type II analysis performs better on average than Type II, based on the results of some simulation studies. Langsrud [39] had a further discussion and found strong reasons to consider Type II analysis as a more appropriate default choice.

Therefore, Type II SS is used for the following analysis. If a significant effect is found, we perform pairwise t-tests as post-hoc analyses. Holm correction [23] is used to adjust the result of pairwise t-tests because Holm corrections is uniformly better than another widely used method: Bonferroni correction, according to the work of Aickin and Gensler [1]. After the analysis of each item, we will discuss the results.

5.1 Anticipation of Danger

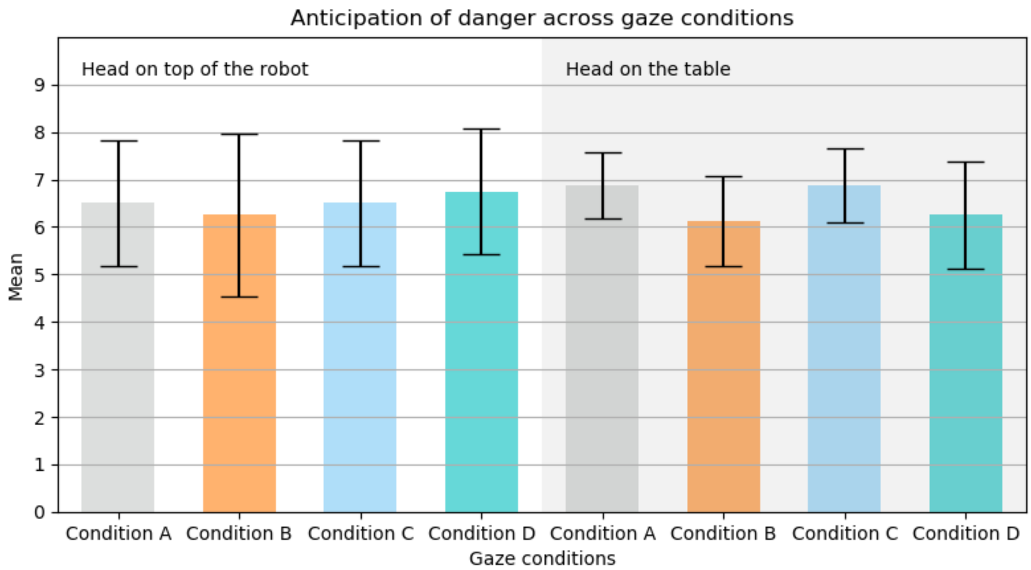

Figure 5.1: Mean scores and standard errors across gaze conditions on anticipation of danger item.

Hypothesis 1 stated that human users would better anticipate danger working with the robot in condition D than in other conditions. To investigate the statistical differences between gaze conditions and robot head positions on anticipation of danger, participants in the two groups responded to the subjective item “I was able to anticipate potential collisions with the robot” between the gaze conditions where they used a nine-point Likert-type scale to give their answers.

The means and standard errors for anticipation of danger are presented in Figure 5.1.

We conducted Mauchly's sphericity test and found that the sphericity is not violated ($X^2(5) = 6.002$, $p = .309$). As shown in Table 5.1, the Levene's test is not significant with a significance level of .05, so equal variances are assumed.

The results of the two-way ANOVA indicated no significant main effect for the robot head position, $F(1, 10) < 0.01$, $p = .983$, partial $\eta^2 < .01$; no significant main effect for gaze condition, $F(3, 30) = 0.45$, $p = .719$, partial $\eta^2 = .04$; and no significant interaction between robot head position and gaze condition, $F(3, 30) = 0.23$, $p = .878$, partial $\eta^2 = .02$. As a result, hypothesis 1 was not supported. The detailed results are shown in Table 5.2 and Table 5.3.

	F	df1	df2	p
A	0.32	1	10	.586
B	0.57	1	10	.469
C	0.18	1	10	.684
D	0.86	1	10	.375

Table 5.1: Test for Equality of Variances (Levene's) for Anticipation of Danger. As we can see, Levene's test is not significant with a significance level of .05.

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Gaze condition	2.896	3	0.965	0.45	.719	.04
Gaze condition * Group	1.448	3	0.483	0.23	.878	.02
Residuals	64.406	30	2.147			

Table 5.2: Within Subjects Effects for Anticipation of Danger. As we can see, the main effect for gaze condition is not significant.

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Group	0.010	1	0.010	< 0.01	.983	$< .01$
Residuals	217.219	10	21.722			

Table 5.3: Between Subjects Effects for Anticipation of Danger. Group means the two different robot head position. The results indicated no significant main effect for robot head position.

The possible reasons for this are that the motion pattern of the robot arm is too simple and its speed is not high enough, so that the participants can completely predict its movement. In addition, participants were asked to perform a practice session consisting of thirty assembly tasks before the start of the experiment. Because in previous pilot studies, we noticed that participants often forgot some task details (e.g., forget to write numbers on paper after completing each assembly task), this practice session was designed to familiarize participants with the tasks to be completed. This seems to exacerbate the phenomenon of participants being too knowledgeable about the behavior of the robot arm's movement. This guess was confirmed by some participants, who gave an open comment "The movement of the robot arm is too simple.". Moreover, some participants gave the feedback "I don't think the robot arm will hurt me." because they knew that the robot arm would stop when there was a potential collision. These reasons together may lead to randomness in the mean scores of the participants on the questionnaire.

5.2 Intent Communication

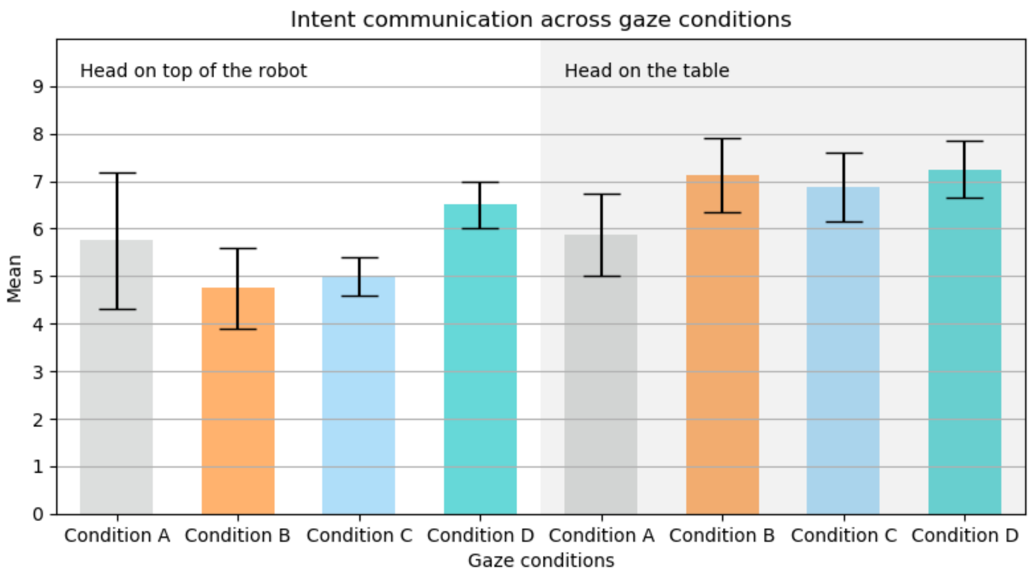

Figure 5.2: Mean scores and standard errors across gaze conditions on intent communication item.

Hypothesis 2 described that robot's gaze behavior in condition D would better communicate the robot's intent to the human user than gaze behaviors in other conditions. To test this hypothesis, participants in two groups responded to the subjective item "The robot communicated its intent clearly" between the gaze conditions where they used a similar nine-point Likert-type scale to give their answers as on anticipation of danger.

The means and standard errors for intent communication are shown in Figure 5.1. Mauchly's sphericity test showed that the sphericity is not violated ($X^2(5) = 7.892$, $p = .165$). In addition, the Levene's test is not significant as shown in Table 5.4, so equal variances are assumed.

We found no significant main effect for robot head position, $F(1, 10) = 2.33$, $p = .158$, partial $\eta^2 = .19$; no significant main effect for gaze condition, $F(3, 30) = 0.94$, $p = .434$, partial $\eta^2 = .09$; and no significant interaction between robot head position and gaze condition, $F(3, 30) = 0.95$, $p = .431$, partial $\eta^2 = .09$. Therefore, hypothesis 2 was not supported. The detailed results are shown in Table 5.5 and Table 5.7.

	F	df1	df2	p
A	0.10	1	10	.764
B	0.02	1	10	.894
C	4.50	1	10	.060
D	0.86	1	10	.376

Table 5.4: Test for Equality of Variances (Levene's) for Intent Communication. As we can see, Levene's test is not significant with a significance level of .05.

Cases	Sum of Squares	df	Mean Square	F	p	η^2_p
Gaze conditions	8.396	3	2.799	0.94	.434	.09
Gaze conditions * Group	8.448	3	2.816	0.95	.431	.09
Residuals	89.406	30	2.980			

Table 5.5: Within Subjects Effects for Intent Communication. As we can see, the main effect for gaze condition is not significant.

Though hypothesis 2 was not supported by the results of ANOVA, we can see a considerable difference in the mean scores for intent communication of two head positions from Figure 5.2. Actually, the mean scores when the robot head is on the table ($M = 6.78$) and when

Table 5.6: Between Subjects Effects for Intent Communication

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Group	17.510	1	17.510	2.33	.158	.19
Residuals	75.219	10	7.522			

Table 5.7: Between Subjects Effects for Intent Communication. Group means the two different robot head position. The results indicated no significant main effect for robot head position.

the robot head is on top of the robot ($M = 5.5$) did give us some feeling of difference. This is also reflected in the p value, $p = .158$, which is already close to the significance level of .05 compared with other p values.

According to the feedback collected from the participants, we found that a large part of the participants was confused about the behavior of the robot arm, giving comments like "The interaction was good but I did not understand why the robot do ... when I got close." or "Sometimes the reaction is random.". One participant directly reveals that he thought he doesn't need to care about the intent of the robot, because he subconsciously believes that robots should cooperate with humans to work, and when a collision occurs, humans should have the priority to continue fetching objects. Furthermore, according to an analysis of recorded video of the experiments, despite the fine-tuning, the robot's head still appeared slow relative to the rapid movements of the human hand. This led to the fact that participants had often finished grasping objects by the time the robot head exhibited obvious behaviors. In this case, the participant has no motivation to guess the robot's intent. So in short, there are two reasons why the robot's head behavior does not improve intent communication much. On the one hand, the movement of the robot's head is not fast enough and in some corner cases not consistent. On the other hand, some participants saw robots as their vassals rather than collaborators.

5.3 Human-perceived Competence of The Robot

Hypothesis 3 stated that gaze behaviors in conditions B, C, and D improve the human-perceived competence of the robot compared with gaze behavior in condition A. To look for statistically significant differences between the robot head positions and gaze conditions

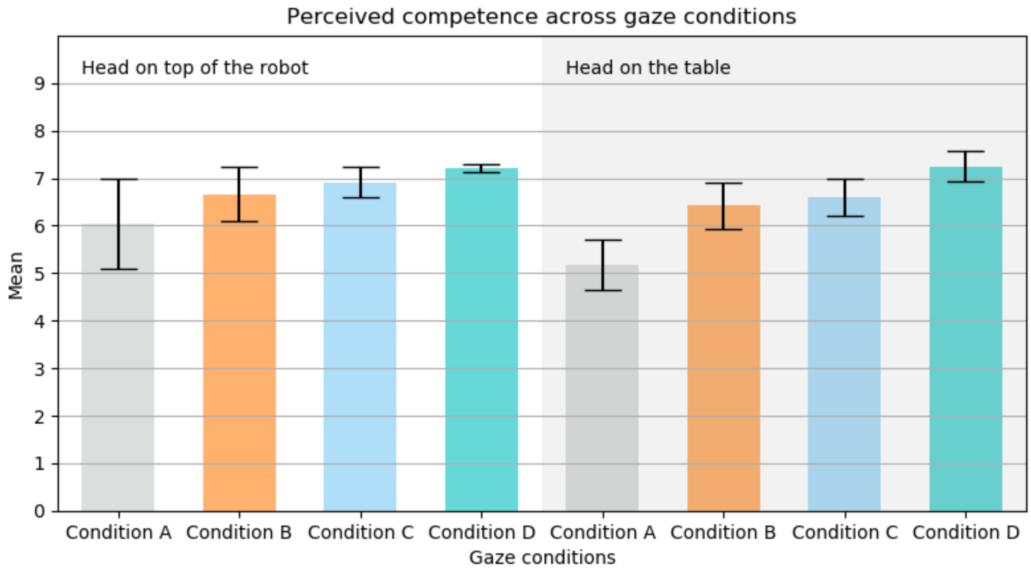

Figure 5.3: Mean scores and standard errors across gaze conditions on human-perceived competence of the robot.

on the human-perceived competence measure, we collected participants' responses to six subjective items ("Knowledgeable", "Interactive", "Responsive", "Capable", "Competent", "Reliable").

Figure 5.3 shows the means and standard errors for human-perceived competence of the robot. Mauchly's sphericity test showed that the sphericity is not violated ($X^2(5) = 8.280$, $p = .144$). However, the Levene's test showed that the homogeneity is violated ($p = .038$), as shown in Table 5.8. Because the JASP software did not provide a correction method for Levenes' test in a two-way ANOVA, and $p = .038$ was not far from a significance level of 0.05, we ignored the correction for Levenes' test.

The results of the ANOVA indicated no significant main effect for the robot head position, $F(1, 10) = 0.56$, $p = .472$, partial $\eta^2 = .05$; a significant main effect for gaze condition, $F(3, 30) = 5.51$, $p = .004$, partial $\eta^2 = .36$; and no significant interaction between the robot head position and gaze condition, $F(3, 30) = 0.32$, $p = .809$, partial $\eta^2 = .03$. The summarized results are shown in Table 5.9 and Table 5.10. Post-hoc testing using pairwise t-tests indicated that human-perceived competence of the robot was significantly higher for gaze condition D than they were for gaze condition A ($p_{holm} = .011$), as shown in

Table 5.11. There was no significant difference between the human-perceived competence of the robot under other gaze conditions, see Table 5.11. Based on the above results, hypothesis 3 was not supported.

	F	df1	df2	p
A	0.07	1	10	.790
B	0.08	1	10	.785
C	1.83	1	10	.206
D	5.72	1	10	.038

Table 5.8: Test for Equality of Variances (Levene's) for Human-perceived Competence of The Robot. As we can see, Levene's test is significant with a significance level of .05.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Gaze conditions	20.012	3	6.671	5.51	.004	.36
Gaze conditions * Group	1.174	3	0.391	0.32	.809	.03
Residuals	36.345	30	1.211			

Table 5.9: Within Subjects Effects for Human-perceived Competence of The Robot. As we can see, the main effect for gaze condition is significant.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Group	1.299	1	1.299	0.56	.472	.05
Residuals	23.281	10	2.328			

Table 5.10: Between Subjects Effects for Human-perceived Competence of The Robot. Group means the two different robot head position. The results indicated no significant main effect for robot head position.

The non-significant between-subjects effects seem to be counter-intuitive because some participants in pilot studies explicitly mentioned that mounting the robot head on top of the robot is too far away. In contrast, some participants in real studies said that indeed they could feel the existence of the head behavior with both robot head positions. This may be due to the natural differences in subjective feelings among different individuals. Another explanation is, when the robot head is on top of the robot, the robot looks more

		Mean Difference	SE	t	Cohen's d	p_{holm}
A	B	-0.937	0.477	-1.967	-0.768	.234
	C	-1.156	0.477	-2.426	-0.947	.107
	D	-1.625	0.477	-3.409	-1.331	.011
B	C	-0.219	0.477	-0.459	-0.179	.667
	D	-0.687	0.477	-1.442	-0.563	.479
C	D	-0.469	0.477	-0.983	-0.384	.667

Table 5.11: Post Hoc Comparisons - Human-perceived Competence of The Robot. A significance level of .05 is used.

human-like, which probably compensates for the disadvantage of the greater distance.

Notably, although we identified significant differences between conditions A and D, some participants reported that head shaking was actually not observed under condition D. Some participants also have the feeling "Out of the corner of my eye I already know the robot head is moving.". Therefore, it is possible that the participants were not sensitive to the details of the different head movements, but merely noticed the sound or vibration when the head switched its behaviors.

5.4 Discomfort

Another subjective item we investigated is discomfort. We used six indicators to collect participants' responses, including "Aggressive", "Awful", "Scary", "Awkward", "Dangerous" and "Strange". Hypothesis 4 was based on this item and proposed that gaze behaviors in conditions B, C, and D would reduce the discomfort of human users compared with gaze behavior in condition A.

The means and standard errors for discomfort are shown in Figure 5.4. Mauchly's sphericity test showed that there is no violation of sphericity ($X^2(5) = 9.467$, $p = .094$). Equal variances are assumed because Levene's test is not significant, as shown in Table 5.12. We found no significant main effect for robot head position, $F(1, 10) = 1.33$, $p = .275$, partial $\eta^2 = .12$; no significant main effect for gaze condition, $F(3, 30) = 0.48$, $p = .699$, partial $\eta^2 = .05$; and no significant interaction between robot head position and gaze condition, $F(3, 30) = 0.88$, $p = .465$, partial $\eta^2 = .08$. The summarized results are shown in Table 5.9 and Table 5.10. Therefore, hypothesis 4 was not supported.

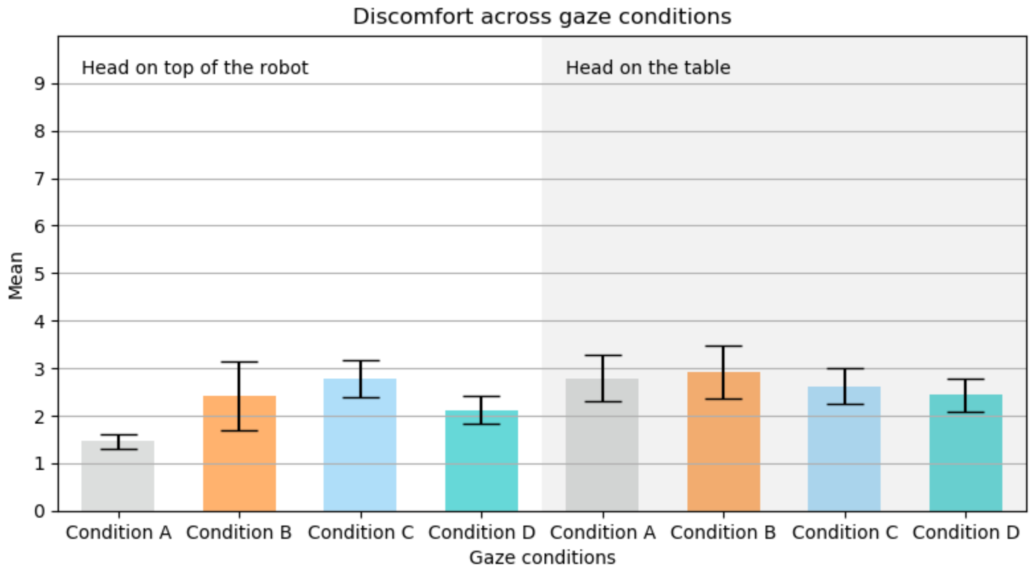

Figure 5.4: Mean scores and standard errors across gaze conditions on discomfort item.

On the item of discomfort, most participants gave a relatively small scale like 1 or 2. This could be due to the cuter facial expression of the robot's head or the participants' belief that the presence of the head is beneficial to them anyway. However, there are also a small number of participants who gave comments such as "The head movement is stiff." and "The head movement is sometimes strange.".

5.5 Task Execution Time

Task execution time represents the total time spent by participants in completing 30 assembly tasks. Due to the low precision of timing, we round up all the values. Hypothesis 5 stated that gaze behavior in conditions D could reduce task execution time.

The means and standard errors for task execution time are shown in Figure 5.5. Mauchly's sphericity test showed that there is no violation of sphericity ($X^2(5) = 6.753$, $p = .242$). Equal variances are assumed because Levene's test is not significant. The results of the ANOVA indicated no significant main effect for the robot head position, $F(1, 10) = 4.31$,

	F	df1	df2	p
A	2.198	1	10	.169
B	0.002	1	10	.968
C	0.693	1	10	.425
D	0.541	1	10	.479

Table 5.12: Test for Equality of Variances (Levene's) for Discomfort. As we can see, Levene's test is not significant with a significance level of .05.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Gaze conditions	1.717	3	0.572	0.48	.699	.05
Gaze conditions * Group	3.129	3	1.04	0.88	.465	.08
Residuals	35.766	30	1.192			

Table 5.13: Within Subjects Effects for Discomfort. As we can see, the main effect for gaze condition is not significant.

$p = .065$, partial $\eta^2 = .30$; no significant main effect for gaze condition, $F(3, 30) = 0.64$, $p = .595$, partial $\eta^2 = .06$; and no significant interaction between robot head position and gaze condition, $F(3, 30) = 0.48$, $p = .700$, partial $\eta^2 = .05$. The summarized results are shown in Table 5.16 and Table 5.17. Therefore, hypothesis 5 was not supported.

Even though not significant, with $p = .065$, there is still a considerable difference between subjects difference in task execution time. Perhaps this is because the increased distance between the robot head and the human allows the human to focus more on the assembly tasks. According to the analysis of the behavior of the participants, most of them were casual about speed control. When encountering a new head behavior, some curious participants even temporarily paused the task in progress to test the behavior of the robot and even play with it. In addition, compared with the robot, the participants have absolute control over the task execution. As the arm approaches, they can decide at will whether to pick up the object faster or let the arm go at first. For these reasons, the presence of the gaze behavior showed little impact on task execution time.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Group	2.611	1	2.611	1.33	.275	.12
Residuals	19.573	10	1.957			

Table 5.14: Between Subjects Effects for Discomfort. Group means the two different robot head position. The results indicated no significant main effect for robot head position.

	F	df1	df2	p
A	2.192	1	10	.169
B	2.187	1	10	.170
C	4.400	1	10	.062
D	1.776	1	10	.212

Table 5.15: Test for Equality of Variances (Levene's) for Task Execution Time. As we can see, Levene's test is not significant with a significance level of .05.

5.6 Familiarity

Familiarity may have a potential impact on task execution time. We investigated the relation between them by measuring a linear correlation. By the ANOVA analysis of task execution time, we already know that there is no significant between-group effect, so we merged the samples of two robot head positions. The results showed that the two variables were weak correlated, $r(10) = .21$, $p = .513$. See Table 5.18. This may be because the collaborative task conducted in the experiment was too easy or the sample size was too small.

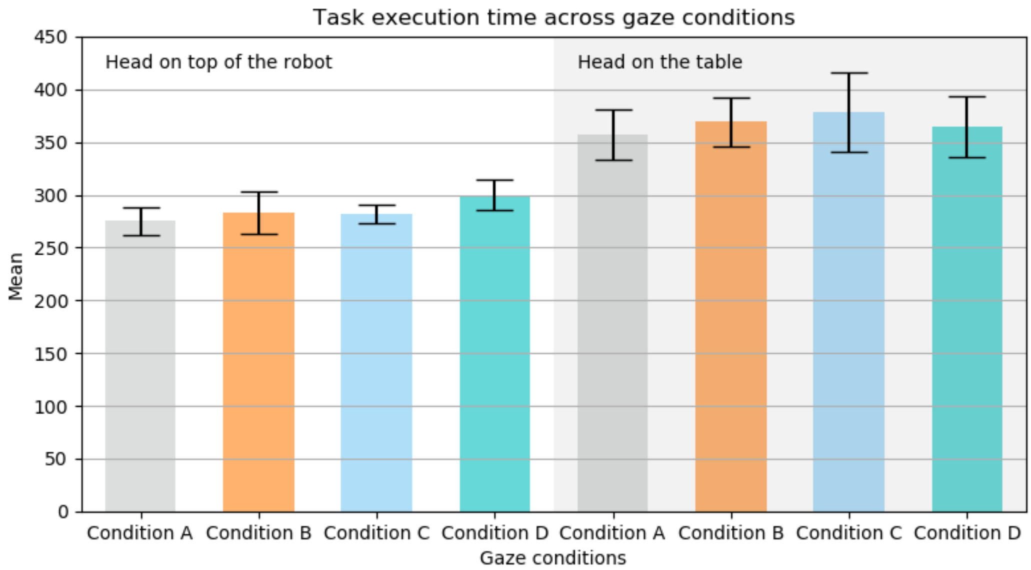

Figure 5.5: Mean and standard errors across gaze conditions on task execution time.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Gaze conditions	1936.250	3	645.417	0.64	.595	.06
Gaze conditions * Group	1442.875	3	480.958	0.48	.700	.05
Residuals	30217.375	30	1007.246			

Table 5.16: Within Subjects Effects of Task Execution Time. As we can see, the main effect for gaze condition is not significant. SC stands for Sphericity Correction, the data is corrected using the Greenhouse-Geisser method.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Group	72270.375	1	72270.375	4.31	.065	.30
Residuals	167864.375	10	16786.437			

Table 5.17: Between Subjects Effects of Task Execution Time. Group means the two different robot head position. The results indicated no significant main effect for robot head position.

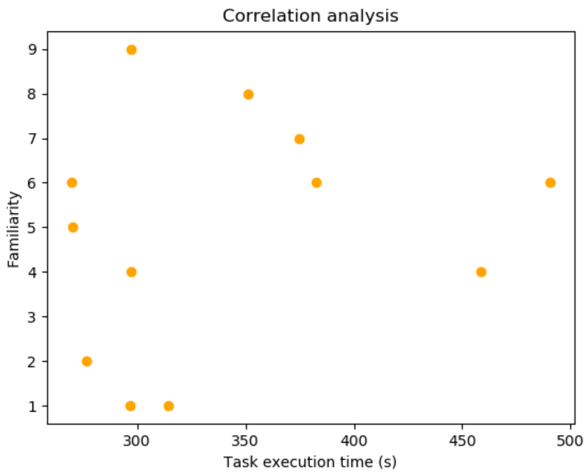

Figure 5.6: The twelve sample points collected. The figure gives us the impression that the correlation between familiarity and task execution time is weak.

Variable		Familiarity	Time
1. Familiarity	Pearson's r	–	
	p-value	–	
2. Time	Pearson's r	0.210	–
	p-value	0.513	–

Table 5.18: Pearson's Correlations of familiarity and task execution time. As we can see, the correlation is weak.

6 Conclusions and Future Work

Based on the results obtained from the experiment, we found that all of the five hypotheses we proposed were not supported. Participants' rating on anticipation of danger, intent communication, human-perceived competence of the robot, discomfort and task execution time were all not influenced by the position of robot head. Our interest in the impact caused by the different locations of the robot head head initially came from the work of Lemasurier et al. [41], which found that the robot signal function is directly impacted by the distance between the signal and the human user. Obviously, the change in robot head position will lead to the change in distance between the robot and the human. The reason we did not get similar results could be that the task performed in the our study was different. However, we still got beneficial insights. An important finding was that the proximity of robot head to human tends to be associated with shorter task execution time, $F(1, 10) = 4.31$, $p = .065$, partial $\eta^2 < .30$. In addition, the gaze behavior had a significant influence on the human-perceived competence of robot. The robot with the gaze behavior in condition D was more competent in human eyes than the robot with gaze behavior in condition A.

Now it is time to answer the two research questions. The first research question is: To what extent, if any, will well-defined robot gaze behavior leads human users to anticipate the robot's intent? The answer is no because the result of the ANOVA showed us that there is no significant main effect for gaze behavior on this item. The second research question is: Do human users get any benefit with a robot equipped with well-defined gaze behavior? The answer is yes. There are two benefits. On the one hand, a well-defined gaze behavior can improve the human-perceived competence of the robot, which may lead to the human being working more confidently, although we did not directly prove this in our study. On the other hand, a well-defined gaze behavior, e.g. with an optimal distance between the robot and the human, has the potential to shorten the task execution time.

For future research, it's a good idea to learn the trajectory from more complex demonstrations. The movement speed of the robot arm also can be further improved. In addition, more vivid facial expressions and smooth head movements have potential to improve the probability of robots being perceived by humans as collaborators. Furthermore, using a questionnaire about comfort instead of discomfort may bring different results, as discussed in section 5.4. Finally, it is also a beneficial attempt to add a questionnaire to study the influence of gaze behavior on human's work confidence.

Bibliography

[1] Mikel Aickin and Helen Gensler. "Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods." In: *American journal of public health* 86.5 (1996), pp. 726–728.

[2] Adrian Bangerter. "Using Pointing and Describing to Achieve Joint Focus of Attention in Dialogue". In: *Psychological science* 15 (July 2004), pp. 415–9. DOI: 10.1111/j.0956-7976.2004.00694.x.

[3] Christoph Bartneck et al. *Human-robot interaction: An introduction*. Cambridge University Press, 2020.

[4] Christoph Bartneck et al. "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots". In: *International journal of social robotics* 1 (2009), pp. 71–81.

[5] Sarah-Jayne Blakemore and Jean Decety. "From the Perception of Action to the Understanding of Intention". In: *Nature reviews. Neuroscience* 2 (Sept. 2001), pp. 561–7. DOI: 10.1038/35086023.

[6] Jean-David Boucher et al. "I Reach Faster When I See You Look: Gaze Effects in Human–Human and Human–Robot Face-to-Face Cooperation". In: *Frontiers in neurorobotics* 6 (May 2012), p. 3. DOI: 10.3389/fnbot.2012.00003.

[7] Colleen M. Carpinella et al. "The Robotic Social Attributes Scale (RoSAS): Development and Validation". In: *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. 2017, pp. 254–262.

[8] Sarah Chambers. *How to Write a Survey Introduction*. https://blog.hubspot.com/service/survey-introduction. Accessed: 2023-05-22.

[9] Fei Chao et al. "A developmental constraint driven approach to developmental robotic hand-eye coordination". In: *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. 2012, pp. 1848–1853. DOI: 10.1109/ROBIO.2012.6491237.

[10] Fei Chao et al. “Enhanced Robotic Hand–Eye Coordination Inspired From Human-Like Behavioral Patterns”. In: *IEEE Transactions on Cognitive and Developmental Systems* 10.2 (2018), pp. 384–396. DOI: 10.1109/TCDS.2016.2620156.

[11] Herbert H. Clark and Susan Brennan. “Grounding in communication”. In: *Perspectives on socially shared cognition*. 1991.

[12] Adrià Colomé and Carme Torras. “Demonstration-free contextualized probabilistic movement primitives, further enhanced with obstacle avoidance”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2017, pp. 3190–3195. DOI: 10.1109/IROS.2017.8206151.

[13] J. Edward, Witaya Wannasuphoprasit, and Michael Peshkin. “Cobots: Robots For Collaboration With Human Operators”. In: (Mar. 1999).

[14] Franz Faul et al. “G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences”. In: *Behavior research methods* 39.2 (2007), pp. 175–191.

[15] Franz Faul et al. “Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses”. In: *Behavior research methods* 41.4 (2009), pp. 1149–1160.

[16] Alessandro Gasparetto et al. “Path planning and trajectory planning algorithms: A general overview”. In: *Motion and operation planning of robotic systems* (2015), pp. 3–27.

[17] F.J. Gravetter and L.A.B. Forzano. *Research Methods for the Behavioral Sciences*. Cengage Learning, 2011. ISBN: 9781111342258. URL: https://books.google.de/books?id=JaIIzgEACAAJ.

[18] Justin Hart et al. “Using Human-Inspired Signals to Disambiguate Navigational Intentions”. In: Nov. 2020, pp. 320–331. ISBN: 978-3-030-62055-4. DOI: 10.1007/978-3-030-62056-1_27.

[19] Justin W. Hart et al. “Gesture, Gaze, Touch, and Hesitation: Timing Cues for Collaborative Work”. In: 2014.

[20] Sayako Hata and Mitsuharu Matsumoto. “Target object announcement combining robot gaze and augmented hand”. In: *2012 Proceedings of SICE Annual Conference (SICE)*. IEEE. 2012, pp. 1669–1672.

[21] David G Herr. “On the history of ANOVA in unbalanced, factorial designs: The first 30 years”. In: *The American Statistician* 40.4 (1986), pp. 265–270.

[22] Guy Hoffman and Xuan Zhao. “A primer for conducting experiments in human–robot interaction”. In: *ACM Transactions on Human-Robot Interaction (THRI)* 10.1 (2020), pp. 1–31.

[23] Sture Holm. “A simple sequentially rejective multiple test procedure”. In: *Scandinavian journal of statistics* (1979), pp. 65–70.

[24] Yong K Hwang and Narendra Ahuja. “Gross motion planning—a survey”. In: *ACM Computing Surveys (CSUR)* 24.3 (1992), pp. 219–291.

[25] ISO Central Secretary. *Robots and robotic devices — Collaborative robots*. en. Standard ISO/TS 15066:2016. Geneva, CH: International Organization for Standardization, 2016. URL: `https://www.iso.org/standard/62996.html`.

[26] ISO Central Secretary. *Robots and robotic devices — Safety requirements for industrial robots — Part 1: Robots*. en. Standard ISO 10218-1:2011. Geneva, CH: International Organization for Standardization, 2011. URL: `https://www.iso.org/standard/51330.html`.

[27] ISO Central Secretary. *Robots and robotic devices — Safety requirements for industrial robots — Part 2: Robot systems and integration*. en. Standard ISO 10218-2:2011. Geneva, CH: International Organization for Standardization, 2011. URL: `https://www.iso.org/standard/41571.html`.

[28] ISO Central Secretary. *Robots and robotic devices — Safety requirements for personal care robots*. en. Standard ISO 13482:2014. Geneva, CH: International Organization for Standardization, 2014. URL: `https://www.iso.org/standard/53820.html`.

[29] JASP Team. *JASP (Version 0.17.1)[Computer software]*. 2023. URL: `https://jasp-stats.org/`.

[30] Long Jin and Yunong Zhang. “G2-Type SRMPC Scheme for Synchronous Manipulation of Two Redundant Robot Arms”. In: *IEEE Transactions on Cybernetics* 45.2 (2015), pp. 153–164. DOI: `10.1109/TCYB.2014.2321390`.

[31] Marcel Adam Just and Patricia A Carpenter. “Eye fixations and cognitive processes”. In: *Cognitive psychology* 8.4 (1976), pp. 441–480.

[32] Adam Kendon. “Some functions of gaze-direction in social interaction”. In: *Acta psychologica* 26 (1967), pp. 22–63.

[33] G. Keppel and T.D. Wickens. *Design and Analysis: A Researcher’s Handbook*. Pearson Prentice Hall, 2004. ISBN: 9780135159415. URL: `https://books.google.de/books?id=SOckAQAAIAAJ`.

[34] Dorothea Koert et al. "Demonstration based trajectory optimization for generalizable robot motions". In: *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. 2016, pp. 515–522. DOI: 10.1109/HUMANOIDS.2016.7803324.

[35] Dorothea Koert et al. "Learning Intention Aware Online Adaptation of Movement Primitives". In: *IEEE Robotics and Automation Letters* 4.4 (2019), pp. 3719–3726. DOI: 10.1109/LRA.2019.2928760.

[36] Nicole Krämer et al. "Theory of Companions: What Can Theoretical Models Contribute to Applications and Understanding of Human-Robot Interaction?" In: *Applied Artificial Intelligence* 25 (July 2011), pp. 474–502. DOI: 10.1080/08839514.2011.587153.

[37] Jörg Krüger, Terje Lien, and Alexander Verl. "Cooperation of human and machines in assembly lines. Keynote paper". In: *CIRP Annals - Manufacturing Technology* 58 (Dec. 2009), pp. 628–646. DOI: 10.1016/j.cirp.2009.09.009.

[38] Alap Kshirsagar et al. "Robot gaze behaviors in human-to-robot handovers". In: *IEEE Robotics and Automation Letters* 5.4 (2020), pp. 6552–6558.

[39] Øyvind Langsrud. "ANOVA for unbalanced data: Use Type II instead of Type III sums of squares". In: *Statistics and Computing* 13.2 (2003), pp. 163–167.

[40] Przemyslaw A Lasota, Terrence Fong, Julie A Shah, et al. "A survey of methods for safe human-robot interaction". In: *Foundations and Trends® in Robotics* 5.4 (2017), pp. 261–349.

[41] Gregory Lemasurier et al. "Methods for expressing robot intent for human–robot collaboration in shared workspaces". In: *ACM Transactions on Human-Robot Interaction (THRI)* 10.4 (2021), pp. 1–27.

[42] James D Lewsey, William P Gardiner, and George Gettinby. "A study of simple unbalanced factorial designs that use type II and type III sums of squares". In: *Communications in Statistics-Simulation and Computation* 26.4 (1997), pp. 1315–1328.

[43] Guilherme J Maeda et al. "Probabilistic movement primitives for coordination of multiple human–robot collaborative tasks". In: *Autonomous Robots* 41.3 (2017), pp. 593–612.

[44] Eloise Matheson et al. "Human–robot collaboration in manufacturing applications: A review". In: *Robotics* 8.4 (2019), p. 100.

[45] *Maximum Likelihood Estimation*. URL: https://online.stat.psu.edu/stat415/lesson/1/1.2 (visited on 12/22/2022).

[46] Margot ME Neggers et al. "Effect of Robot Gazing Behavior on Human Comfort and Robot Predictability in Navigation". In: *2022 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO)*. IEEE. 2022, pp. 1–6.

[47] Stephanie Olson, Moaed Abd, and Erik D Engeberg. "Human-Inspired Robotic Eye-Hand Coordination Enables New Communication Channels Between Humans and Robots". In: *International Journal of Social Robotics* 13.5 (2021), pp. 1033–1046.

[48] Omron. *Kollaborative Roboter*. `https://industrial.omron.de/de/products/collaborative-robots`. Accessed: 2022-12-07.

[49] Alex Owen-Hill. *What You Need to Know About Collaborative Robot Safety*. `https://robodk.com/blog/collaborative-robot-safety`. Accessed: 2022-12-07.

[50] Alexandros Paraschos et al. "Probabilistic movement primitives". In: *Advances in neural information processing systems* 26 (2013).

[51] COMMENT PICKER. *RANDOM NUMBER GENERATOR*. 2023. URL: `https://commentpicker.com/random-number-generator.php`.

[52] Peter Poór, Tomáš Broum, and Josef Basl. "Role of Collaborative Robots in Industry 4.0 with Target on Education in Industrial Engineering". In: *2019 4th International Conference on Control, Robotics and Cybernetics (CRC)*. 2019, pp. 42–46. DOI: `10.1109/CRC.2019.00018`.

[53] David Robinson. "Manwatching: A Field Guide to Human Behaviour. By Desmond Morris. London: Jonathan Cape. 1977. Pp 320. £7.95." In: *British Journal of Psychiatry* 132.3 (1978), pp. 313–314. DOI: `10.1192/S0007125000283505`.

[54] Richard J Rossi. *Mathematical statistics: an introduction to likelihood based inference*. John Wiley & Sons, 2018.

[55] Stefan Schaal. "Learning from demonstration". In: *Advances in neural information processing systems* 9 (1996).

[56] Fahad Sherwani, Muhammad Mujtaba Asad, and Babul Salam Kader K Ibrahim. "Collaborative robots and industrial revolution 4.0 (ir 4.0)". In: *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*. IEEE. 2020, pp. 1–5.

[57] Moondeep C. Shrestha et al. "Communicating Directional Intent in Robot Navigation using Projection Indicators". In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 2018, pp. 746–751. DOI: `10.1109/ROMAN.2018.8525528`.

[58] F Michael Speed, Ronald R Hocking, and OiP Hackney. “Methods of analysis of linear models with unbalanced data”. In: *Journal of the American Statistical Association* 73.361 (1978), pp. 105–112.

[59] Michael Tomasello et al. “Understanding and sharing intentions: The origins of cultural cognition”. In: *Behavioral and brain sciences* 28.5 (2005), pp. 675–691.

[60] Federico Vicentini. “Collaborative robotics: a survey”. In: *Journal of Mechanical Design* 143.4 (2021).

[61] Wikipedia contributors. *Maximum likelihood estimation*. [Online; accessed 22-December-2022]. 2022. URL: https://en.wikipedia.org/wiki/Maximum_likelihood_estimation.

[62] Wikipedia contributors. *Naive Bayes classifier*. [Online; accessed 23-November-2022]. 2022. URL: https://en.wikipedia.org/wiki/Naive_Bayes_classifier.

[63] Wikipedia contributors. *Quadratic programming*. [Online; accessed 25-January-2023]. 2023. URL: https://en.wikipedia.org/wiki/Quadratic_programming.

[64] Wikipedia contributors. *Radial basis function*. [Online; accessed 22-December-2022]. 2022. URL: https://en.wikipedia.org/wiki/Radial_basis_function.

[65] Min Zhang et al. “Dynamic artificial potential field based multi-robot formation control”. In: *2010 IEEE Instrumentation & Measurement Technology Conference Proceedings*. IEEE. 2010, pp. 1530–1534.

[66] Tao Zhou and Bertram E. Shi. “Learning Visuomotor Transformations and End Effector Appearance by Local Visual Consistency”. In: *IEEE Transactions on Cognitive and Developmental Systems* 8.1 (2016), pp. 60–69. DOI: 10.1109/TAMD.2015.2502758.