COMENIUS UNIVERSITY BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

ASSESSING THE LEGIBILITY OF HUMANOID ROBOT NICO MOVEMENT

Master's thesis

2025

Hana Hornáčková

COMENIUS UNIVERSITY BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

ASSESSING THE LEGIBILITY OF HUMANOID ROBOT NICO MOVEMENT

Master's thesis

Study programme: Cognitive Science

Field of study: Applied Informatics

Supervising department: Department of Applied Informatics

Supervisor: Andrej Lúčny, RNDr.

Bratislava 2025

Hana Hornáčková

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname: Hana Hornáčková
Study programme: Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study: Computer Science
Type of Thesis: Diploma Thesis
Language of Thesis: English
Secondary language: Slovak

Title: Assessing the Legibility of Humanoid Robot NICO Movement

Annotation: Smooth human–robot interaction (HRI) scenarios must involve robots supporting humans' ability to interpret, predict, and feel safe around robotic actions. One of the critical features is the legibility of the robot movement trajectory characterized by its distinctiveness that helps the observer to disambiguate the robot's intent. Humanoid robot NICO is a relatively new, affordable platform designed for HRI experiments whose potential has yet to be explored.

Aim: Investigate the concept of legibility, based on methods of generating repeatable robotic arm movements. Using both behavioral responses and the questionnaires of participants, analyze the results of an HRI experiment, where the robot is aiming to touch the horizontal screen in various pointing-gaze conditions.

Literature:
Dragan A.D. , Lee K. .T., Srinivasa S.S. (2013) Legibility and predictability of Robot Motion. In 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI).
Kerzel M., Strahl E., Magg, S., Navarro-Guerrero N., Heinrich S., Wermter, S. (2017) NICO - Neuro-Inspired COmpanion: A developmental humanoid robot platform for multimodal interaction. In: IEEE RO-MAN. pp. 113–120.
Sciutti, A., Sandini, G.: Interacting with robots to investigate the bases of social interaction. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25(12), 2295–2304

Supervisor: RNDr. Andrej Lúčny, PhD.
Department: FMFI.KAI - Department of Applied Informatics
Head of department: doc. RNDr. Tatiana Jajcayová, PhD.

Assigned: 19.03.2024

Approved: 24.03.2024 prof. Ing. Igor Farkaš, Dr.
Guarantor of Study Programme

Student

Supervisor

Declaration of Originality

I hereby declare that this master's thesis is my own original work and that I have used only the sources and literature listed in the bibliography. All sources and quotations used have been properly cited.

……

signature

Acknowledgment

I would like to express my sincere gratitude to my supervisor, RNDr. Andrej Lúčny, for his invaluable guidance, support, and encouragement throughout the course of my master's thesis. His expertise, insightful feedback, and patience were instrumental in shaping my research and helping me overcome various challenges. I greatly appreciate the time and effort he devoted to my work, as well as his willingness to share knowledge and provide constructive advice.

Abstrakt

Efektívna komunikácia zámerov robota (čitateľnosť) je kľúčová pre úspešnú interakciu človek-robot (HRI). Táto štúdia skúma, ako vzájomné pôsobenie pohľadu, giest ukazovania, dĺžky trajektórie a postojov pozorovateľov ovplyvňuje dekódovanie zámerov. Pomocou robota NICO sme testovali podmienky s rôznymi modalitami podnetov (iba pohľad, iba ukazovanie, multimodálne) a mierou skrátenia trajektórie (60 % vs. 80 % dokončenia). Výsledky ukazujú, že multimodálne podnety (pohľad + ukazovanie) zvyšujú presnosť vďaka integrácii vizuálno-priestorových a motorických signálov. Predĺžené trajektórie (80 %) zlepšujú presnosť pri ukazovaní a multimodálnych pokusoch, čím sa redukuje neistota extrapolácie, zatiaľ čo pohľad skracuje reakčný čas. Zhodné podmienky sú spojené s vyššou presnosťou oproti nezhodným. Predošlé postoje k robotom preukázali len zanedbateľnú koreláciu s výkonom. Tieto poznatky posúvajú vpred dizajn HRI s dôrazom na ľudské percepčné heuristiky prostredníctvom adaptívneho, spoločensky transparentného pohybu.

Kľúčové slová: interakcia robot-človek, robot, pohyb, čitateľnosť, umelá inteligencia

Abstract

Effective communication of robotic intent (*legibility*) is critical for human-robot interaction (HRI). This study examines how gaze, pointing cues, trajectory duration, and observer attitudes shape intention decoding. Using the robot NICO, we tested conditions varying in cue modality (gaze-only, pointing-only, multimodal) and truncation (60% vs. 80% completion). Results show multimodal cues (gaze + pointing) enhance accuracy by integrating visuospatial and motoric signals. Extended trajectories (80%) improve accuracy in pointing and multimodal trials, reducing extrapolation uncertainty, while gaze expedites reaction times. Congruent conditions are linked to better accuracy scores than incongruent. Pre-existing robot attitudes show negligible correlation with performance. This advances HRI design, prioritising human perceptual heuristics through adaptive, socially transparent motion.

Keywords: human-robot interaction, robot, movement, legibility, artificial intelligence

TABLE OF CONTENTS

Introduction

We live in an era where robotics is constantly advancing and the subsequent interaction between humans and robots is becoming increasingly diverse and nuanced. Robots have long ceased to be exclusively tools on production lines but are increasingly being deployed in various areas such as services, healthcare, logistics and the home environment. Currently, robots are increasingly used in the areas of hospitality, services, warehousing, transport and in households. At the same time, however, this trend creates an inevitable need for human connection with robotic systems. If a robotic system cannot clearly communicate its intentions, misunderstandings may occur, cooperation efficiency may decrease, or user safety may be compromised. However, one of the characteristics that supports such interaction is the robot's ability to convey intent through movement, which is the subject of a concept known as legibility, which is analysed in detail in the professional literature.

Specifically, legibility is defined as “the extent to which a human observer can correctly and quickly determine a robot’s goal based on observed movement that is incomplete.” Legible movement therefore does not result from the ultimate success or accuracy of the movement, but from its communicative prowess in how well (and effectively) the robot can convey what it is attempting through its movement. The importance of legible movement is due to the fact that humans naturally interpret meaning from the movements of agents in their environment that appear to be goal-directed. Therefore, if robots can confidently adapt this behaviour, humans will trust its goal and feel safe in its presence—in turn, humans will believe that they can successfully collaborate in robotic mode. Legible movement is particularly important in situations where a human and a robot share a common workspace—for example, when performing object manipulation while trying to avoid collision or manipulation through object transfer. This thesis attempts to address the concept of legibility of robot movement through a theoretical and experimental approach. We will analyse the technical terminology, available literature, and research developments in the field of legibility of movement.

Furthermore, legibility will be separated from prediction, a formal measurement model will be discussed, as well as peripheral influences such as observer-determined fields of view or eye trajectories/fixation movements. The second half will take this theoretical

learning and apply it to a real experiment with a social robot named NICO to see if gesture and gaze of its hand affect the observer's ability to successfully predict a target.

The aim of this thesis is to offer a comprehensive overview of the issue of legibility of movement in robotics in relation to the world of robotics, as well as to confirm through experimental control that human understanding of robot intentions is not only natural, but also testable. In the theoretical part, we will therefore rely on current professional literature and proven models that form the basis for developmentally and socially successful cognitively legible robotic systems. The work contributes to deepening knowledge in the field of legibility of movement and formulates recommendations for future research in the context of experimental robotics.

1. Theoretical Framework

1.1 Legibility vs. Predictability

The ability of a robot to express what it intends to do is of great importance in the field of human-robot interaction. The robot's motion is often the only source of information from which a human can infer what the robot is planning or trying to do. However, in this field, there are two concepts that are often combined but are very different: legibility and predictability. They concern the understanding of the robot's motion, but they emerge from opposite inferential starting positions and have different implications for trajectory generation (Dragan, Lee & Srinivasa, 2013).

Specifically, a predictable action is one that is consistent with human expectations of the goal. This means that if the observer is informed about the robot's goal, he or she should be able to determine the trajectory that the robot will take to achieve such an action. This type of movement occurs when the trajectory is most optimized according to a cost function of some efficiency – whether that means over the shortest distance or the shortest time to complete. Furthermore, it is based on the so-called principle of rational action, which assumes that agents (including robots) act efficiently and purposefully (Henry, n. d.; Dragan & Srinivasa, 2014).

In contrast, readability stems from how well the observed trajectory of the robot can have its goal determined. It is not the trajectory that we would expect given a known goal, but how well we can estimate what goal the robot is pursuing based on the movement. Therefore, readable action helps the observer to accurately and quickly guess the robot's intention – often before its trajectory is completed. Formally, this is represented by Bayesian inference, where an incomplete trajectory leads to what is the most likely goal (Dragan & Srinivasa, 2013).

While these definitions may seem similar on the surface—both relate to human understanding of robot movement—they are actively at odds. The most predictable (i.e., efficient) action does not mean it is legible; in fact, the opposite may be true. For example, if there are many targets in close proximity, a robot may intentionally choose a strangely eccentric trajectory to help facilitate legibility—even though it might be more efficient (Dragan, Lee, & Srinivasa, 2013).

Ultimately, the rationale for these two approaches lies in the cognitive model of inference. Predictability is based on the assumption that a human knows the target and can therefore infer the likely intended path. Legibility exists in direct opposition—to inference through movement as an attempt to determine its target. These inference approaches differ, therefore even trajectories optimized for readability can differ significantly from those optimized for predictability (Dragan & Srinivasa, 2013).

Interestingly, from a psychological perspective, both concepts are assessed by teleological inference, i.e. the human tendency to assume that the actions of others were intentional. This has been repeatedly demonstrated in children and adults, compared to the assumption of intentionality when a given action appears to be purposeful (Tomasello et al., 2005; Csibra & Gergely, 2007).

This principle is therefore also relevant for inferences related to the legibility of robotic movement. For example, when assessing legibility of action, the probability that other perceivers correctly interpret the intention from the intentional action is taken into account. This is formally expressed as the assessment of the posterior probability of intention conditional on the ongoing action via Bayes' theorem. The resulting legibility is defined as the accumulated probability over time, with emphasis on the parts of movements that occur earlier in the temporal sequence, as they allow the observer to process information about the robot's intention more quickly (Nikolaidis, Dragan & Srinivasa, 2016; Busch et al., 2017).

In reality, however, movement efficiency and legibility of movement may compete with each other. For example, if a robot needs to reach an object that is on a table, it should do so as quickly as possible. It can take the path of least resistance in a straight line to the object, which minimizes the time and energy spent; however, this is an incomprehensible action if two objects are actually pressed against each other on the table. Therefore, to maximize the legibility of intention, the robot should raise its hand above the object and then point it towards the target, thereby clearly indicating its intention (Dragan & Srinivasa, 2013).

Since real-world environments are complex and diverse, there is an effort to create models that also take into account changing conditions and observer expectations. Many of them use the functional gradient method to obtain a trajectory in relation to maximum readability and efficiency over time and space. There are even models that relate perspective

– the position and distance from where the observer sees the agent's action – to create so-called viewpoint-based readability (Nikolaidis et al., 2016).

These findings have direct implications for the design of robotic systems intended to interact with humans. In human-robot interactions within shared environments, operation needs to be more than efficient; it also needs to be understandable. This creates a conflict between predictability and readability or choosing one over the other based on the situational parameters of the task. If the environment is industrial and can tolerate less understandable actions, where predictability can increase efficiency, then yes. If it is in a home environment or a social robotics environment, then understanding actions and safety during interaction must come first (Alami et al., 2006; Hahn & Stone, 2021).

1.2 Legibility of Motion in Robotics

Since the beginning of the second decade of the 21st century, the literature on human-robot interaction has increasingly emphasized the importance of legibility of movement, with research moving from purely performance indicators towards aspects affecting the social acceptance of robots. The first systematic efforts to grasp the legibility of robotic movement as a separate concept appeared in research focused on trajectory planning, which, in addition to efficiency, was intended to enable an unambiguous understanding of the robot's intention by the observer. This meant that the robot's trajectory had to be planned taking into account not only kinematic inferences, but also psychological inferences resulting from the human understanding of the robot's movements (Dragan, Lee & Srinivasa, 2013; Dragan & Srinivasa, 2014).

Early research in this area drew on cognitive psychology, in particular teleological reasoning, according to which people naturally interpret the behaviour of others as purposeful and intentional (Tomasello et al., 2005). Such findings were then translated into robotic systems under the assumption that the observer will attribute action to the robot's movements that result from an innate intention. Therefore, it is essential that the trajectory is not only kinematically accurate but also communicatively efficient. These early approaches led to the design of analytical models that allowed the generation of movements that increase the probability of correctly recognizing the robot's intention during task execution. The logic of these models was based on Bayesian reasoning, where a trajectory

given to an observer infers a probable intention (Dragan & Srinivasa, 2013; Nikolaidis, Dragan & Srinivasa, 2016).

One of the first attempts to quantify legibility as a construct occurred by considering legibility as an optimization problem, in which the input is a uniform probability distribution of potential targets, and the output is the optimization of the informativeness of movements associated with the correct target. These models were implemented using functional gradient methods, which allowed for the generation of trajectories that deviate from canonical, rational, or predictable movements because they favour greater resolution between potential targets (Dragan & Srinivasa, 2013).

For example, instead of creating the shortest trajectory to target A, the robot can gesture over what appears to be target A, or create trajectories that avoid ambiguity with potential, distinguishable trajectories associated with other targets in the vicinity. Later research extended this issue in new directions, such as adaptive models that additionally took into account the relative position and viewpoint of the observer (Nikolaidis et al., 2016).

In terms of metric evaluation, legibility began to be quantified through probabilistic measures that assessed whether and how likely it was that an observer would correctly identify the robot's target before the end of its trajectory.

This was later extended to time measurements, as readability is more than accuracy, but also speed. A trajectory that allows for rapid detection of intent is more readable than a trajectory that takes longer to process or leads to ambiguity (Busch et al., 2017). Studies have emerged using implicit behavioural determinants such as reaction time, guessing success, or observers' eye movements as indicators of readability (Wallkötter, Chetouani & Castellano, 2022).

In recent years, the use of machine learning and generative models to generate legible movements has also expanded. These models are trained on the basis of human feedback and demonstrations and can thus create trajectories that are not only physically feasible but also socially understandable. Techniques such as behavioural cloning, generative models, or reinforcement learning with a penalty for ambiguous trajectories are used here (Bronars & Xu, 2023; Florence, Manuelli & Tedrake, 2022).

An important aspect is also the robustness of such models in changing environments, variability of goals, or ambiguous situations where it is necessary to adjust the movement in real time in order to maintain intelligibility. The relationship between the trajectory and the robot's intention is currently considered the core of legibility. The trajectory is understood as a medium of communication in which the movement is the carrier of information about the goal. From the perspective of robotic behaviour design, this means that robots must be able to plan and execute movements not only with regard to the physical execution of the task, but also to how these movements will be interpreted by humans. For this reason, legibility becomes a key property, especially in areas where robots share space with humans, such as healthcare, logistics, assisted mobility or domestic environments (Alami et al., 2006; Hahn & Stone, 2021).

The link between robot movement and intention is therefore not only a question of navigation, but above all a question of transparency of intention. The challenge for designers of such systems is to ensure that every aspect of the movement – from speed to direction to trajectory – contributes to the creation of a mental model that allows the observer to correctly and safely understand what the robot wants to do. For this reason, legibility is considered an integral part of the modern development of cognitive robotics and its application in real-world social environments.

1.3 Modelling Legibility

Modelling legible robot motion is an engineering approach to designing robotic systems that are easily understood by a human observer. Standard motion planning algorithms typically optimize trajectories in terms of execution success, stability, and safety. When legibility—the ability of a motion to convey the robot’s intent to a human observer—is also included in trajectory planning, the trajectory design process becomes more complex. Modern research focuses on formalizing legibility using probabilistic models, specifically the Bayesian approach, which estimates the probability that a human will correctly identify

a robot's target based on the observed trajectory. (Dragan & Srinivasa, 2013; Nikolaidis, Dragan & Srinivasa, 2016).

The fundamental premise of Bayesian modelling of legibility is the idea that a human observer, when watching a robot's trajectory, tries to infer which goal the robot is heading toward. This inference occurs even during the observation of an incomplete trajectory, that is, before the robot reaches its goal. The Bayesian approach provides a mathematical framework for formalizing this process. The calculation is performed using Bayes' rule, which describes how the probability of a goal is updated based on new information—in this case, a fragment of the trajectory. Formally, the posterior probability of a goal G given a trajectory ξ is calculated as: $P(G|\xi) = \frac{P(\xi|G)P(G)}{P(\xi)}$,

where $P(\xi|G)$ represents the likelihood of the trajectory given a known goal, $P(G)$ is the prior probability of the goal, and $P(\xi)$ is the probability of the trajectory independent of the goal (Dragan & Srinivasa, 2013).

This model allows us to quantify how well a given trajectory communicates the robot's intention: a trajectory is legible if it maximizes the observer's probability of selecting the correct target. This means that we need to create a so-called legibility score about how informative a given trajectory is with respect to its intention. In the simplest model, this is the maximum posterior probability of the correct target at a given time. In more complex models, this is summed over the trajectory with a higher weight assigned to the most informative points at the beginning of the action (Nikolaidis, Dragan & Srinivasa, 2016; Busch et al., 2017).

The Bayesian approach is flexible and efficient, but its reliability depends on the quality of the probabilities and assumptions used. In later years, more and more models began to appear that combined this with cost functions. Cost functions are essentially the "costs" or penalties for executing a particular trajectory in terms of legibility. Cost functions penalize trajectories that are ambiguous or difficult to interpret, for example, if they lead to multiple targets with similar probabilities or if they take longer to understand. The goal of optimization methods is to find a trajectory that minimizes cost and maximizes readability (Dragan, Lee & Srinivasa, 2013).

The most discussed tool for readability optimization is the functional gradient – a function that allows you to gradually change the shape of the trajectory so that it is more informative to the observer. The functional gradient approach examines small variations in the trajectory and evaluates how they affect the probability of correctly recognizing the target. This means that some trajectories may be chosen to be more readable, even though they are not the shortest or most efficient according to traditional criteria (Dragan & Srinivasa, 2013; Nikolaidis et al., 2016).

Another important factor in modelling for readability is the assumed a priori distribution of intentions. For example, if it is likely that the robot can move towards three different goals, but one occurs more frequently or is more critical than the others, you should adjust the calculation accordingly. This means that in more complex implementations, the dynamic updating of the prior probabilities based on the observer's contextual information, previous actions, or robot preferences is also taken into account (Wallkötter, Chetouani & Castellano, 2022).

When evaluating readability scores, the so-called trajectory informativeness is sometimes used as the difference between the posterior probabilities of the goals. This allows us to assess how effectively a trajectory distinguishes one goal from the others – does the observer have enough reasons to eliminate false possibilities? It can be understood as a measure of the contrast between the goals – the greater the contrast, the more readable it is (Nikolaidis, Dragan & Srinivasa, 2016; Busch et al., 2017).

Theoretically, these measures apply to trajectory planning within robotic arms, other autonomous vehicles and drones, or social robots in the home. However, input differences between observers also have a major impact on readability ratings. Therefore, some applied authors include variables such as reaction times, target identification accuracy, saccade trajectories, or even neurophysiological data collected during passive observation of robotic movements. These behavioural factors are then used retrospectively to adjust the optimization functions. This allows current readability models to better adapt to the type of user or the context of the situation, which generally contributes to an increasingly effective robustness of their performance in real life (Busch et al., 2017; Wallkötter, Chetouani & Castellano, 2022).

Recently, motion legibility has been extended through deep learning, which is trained through generative models such as variational autoencoders and latent diffusion models to

distinguish legible motion from a human perspective (Bronars & Xu, 2023; Florence, Manuelli & Tedrake, 2022). Such processes avoid the need for manual posterior probability computation by learning distributions from the training data. The advantage of such a process is that the trained model can automatically distinguish what is legible and useful through a trajectory to a human agent without the need for a manually specified cost function.

Furthermore, using reinforcement learning methods, a robot can learn useful legible motion through engagement with the environment and human feedback (Zhao et al., 2020; Ravichandar et al., 2020). Therefore, Bayesian developments, cost function improvements, and the findings of newer generations of machine learning provide new expectations regarding legibility, where robots can not only function, but also appear to function and are easy to understand. This is a fundamental prerequisite for the development of cognitive robots, whose behaviour should clearly communicate intent and promote trust, safety, and efficiency in human-autonomy interactions. Legibility thus exists as a socio-technical characteristic of robotic locomotion in natural environments.

1.4 The impact of gaze and smooth observation on readability

The perception of robot motion plays a key role in the process of intention inference. How a person understands a robot's intention is not determined solely by the shape or purpose of the trajectory, but also by the perspective from which the trajectory is observed. Therefore, the literature increasingly emphasizes the so-called visual perspective of the observer as an important parameter affecting the legibility of the movement. In this context, Nikolaidis, Dragan, and Srinivasa (2016) introduced the concept of viewpoint-based legibility, which directly involves the observer's position and angle of view in the process of planning a robot trajectory. Their models showed that the same movement can be interpreted as unambiguous or ambiguous depending on the location and angle from which the person observes it.

A trajectory optimized regardless of the human position may thus fail to communicate the intention, while another, less efficient in terms of performance, may be significantly more legible if designed with the observer's visual field in mind. Further research by the same authors confirms that legibility is a spatially conditioned phenomenon. Their experiments showed that if a person is located outside the main axis of movement, the probability of correctly determining the robot's goal before its trajectory ends decreases

(Nikolaidis et al., 2016). This knowledge implies that legibility is not a universal property of the trajectory, but must always be assessed in the context of a specific spatial arrangement. This framework is followed by Wallkötter, Chetouani and Castellano (2022), who expanded the understanding of the influence of visual access to the issue of movement occlusion. Their concept of occlusion-based legibility points out that if part of the trajectory is physically obscured – for example, by another object, part of the robot's body or another person – there is a significant decrease in legibility. In their experiments, they measured how the success of determining the goal changes with different types of occlusions.

The results showed that even short-term or partial visual blockages at key moments of movement can negatively affect the observer's ability to correctly interpret the robot's intention.

These findings are particularly relevant in environments where dynamic human-robot interaction is expected, such as hospitals, schools, domestic spaces, or manufacturing halls. Dragan and Srinivasa (2013) have already pointed out in earlier work that effective trajectory planning must also take into account which parts of the movement are most informative to the human. From a visual processing perspective, these are usually the parts that are at the beginning of the trajectory and are also fully visible. If these segments are obscured, the human's ability to correctly infer intention decreases dramatically. The authors therefore proposed motion planning with redundancy - that is, some movements are repeated or overemphasized to ensure their intelligibility even in the event of partial visual interruption.

Moreover, recent research points to the need to consider multiple observers simultaneously. In real-world environments, it is often the case that a robot interacts with multiple people at the same time, each with a different perspective, a different location in space, and a different cognitive framework. Nikolaidis et al. (2016) therefore introduced the idea of multi-view optimization, in which readability is evaluated as an aggregate function across different perspectives. In this case, the trajectory is optimized to be as readable as possible for the largest possible number of observers. This approach requires the integration of a human-oriented sensor (e.g., cameras or depth sensors) that monitors where people are and how well they can see key parts of the robot's movement. From a practical perspective, this knowledge means that when designing trajectories, it is also necessary to consider the so-called visual accessibility.

This means that planning should not only take place in the physical space of movement, but also in the so-called visual space, where what a person will actually see and process is taken into account. If the goal is to achieve legible behaviour, then the trajectory must be designed so that key information is accessible, visible and unambiguous from as many perspectives as possible. This can be done by using, for example, body expressions, a deliberate gesture towards the goal, or significant movements that are also readable from peripheral vision. Some research, such as the work of Busch, Mörtl and Hirche (2017), also shows that people tend to follow the parts of the movement that are either the fastest or the most unusual the most. This is why it is recommended to highlight the beginning of the movement, or create a so-called "information window" in the first stages of the trajectory, where the robot will provide as many signals about its intention as possible. This knowledge is also directly applied to the creation of an experimental design in the practical part of this work, where it will be investigated how different combinations of view and direction of movement will affect a person's ability to identify a goal in a social context.

The results of all the above research clearly show that legibility is not only a question of what the robot does, but also of where we observe it from. In the context of modern social robotic systems, it is therefore essential to approach trajectory design as an information-rich process that takes into account the dynamics of the environment, the visual abilities of observers, and potential visual limitations. Without this, the robot could behave technically correctly but at the same time be communicatively illegible – which is a fundamental problem in an environment with human interaction.

1.5 Readability modelling: simple approaches

While the ongoing debate about the legibility of robotic motion suggests a growing consensus on the potential of both complex generative models and simpler alternatives to achieve comparable results, a less complicated solution is not only more feasible in terms of computation time and explainability. Simpler alternatives also exist that suggest generalization of concepts, such as finding an appropriate linear perceptron classification for predicting the goal of robot movement. A perceptron is a binary classifier that distinguishes between two classes using a linear combination of input features. The use of a perceptron as a classifier would successfully apply in such a way that the features consist of particular

points along the trajectory or derived aspects of motion (e.g. direction, speed, time point) and the classification is final movement goal.

This is particularly useful in scenarios with a limited number of possible outcomes, such as with the work by Busch, Mörtl, and Hirche (2017), who found that participants were able to infer the robot's goal near the end of the non-linear trajectory by observing their reactions throughout the movement. Therefore, using a perceptron to predict legibility not only facilitates the utilization of specific points along motion trajectories but also allows researchers to better understand which impulses were most useful to predict intention recognition through weight allocation. This means that the training set can give weights to specific sections of motion, allowing researchers to determine which orientations render motion intention recognition most effective and may be clearer if an observer needs to provide clear visual distinction between very similar motions or if only partial segments of the overall motion plan are visible to the observer. Furthermore, as it can be embedded into any trial without the need for modification or extensive manipulation from a preexisting experimental design to merely assess legibility in real or quasi-realistic settings, it serves as an effective method for evaluating legibility in practical scenarios.

Yet while limited in computational ability, it is interpretable, both as a foundation for a more complex model of legibility classification or for post-hoc interpretation of the findings. Another noteworthy method involves using gradient descent as an optimization technique to render trajectories more legible.

Where classical motion planners seek to minimize cost in terms of efficiency (time, energy, distance), the legibility planner seeks to maximize the likelihood of an observer correctly recognizing its intent. By iteratively modifying the trajectory's shape, one can converge to all forms that yield an increased posterior probability of the correct intention.

Dragan and Srinivas (2013) explore this notion extensively; they redefine legibility as an optimization problem and demonstrate that gradient-based approaches can increase the differences between posteriors of individual intentions. Ultimately, the resulting trajectory may be less efficient from a classically rational perspective, but much more informative to the observer.

Essentially, gradient descent allows for the creation of trajectories that purposefully avoid the fastest route to reduce uncertainty. Imagine a robot with multiple targets all in a similar proximity. Where a typical planner would generate an average or straight trajectory that maybe an onlooker mistakes—using gradient descent allows for the movement to be explicitly focused on just one target from the outset even if that means greater energy

expenditure to get to such a conclusion. Such optimization works well in dynamic environments where trajectories are constantly shifting, and responders must rely on the expected actions of the robot. In fact, using gradient descent becomes a useful tool for dynamically adjusting trajectories in real-time.

The final technique that's growing popularity within the realm of readable motion is called motion blending. Motion blending occurs when multiple components of movement or gestures are blended together to create a trajectory that better expresses what the robot is doing than what any one component could do on its own. Therefore, it becomes two or more partial trajectories or gestures which work together to provide better informative power. Blending is based on the principle that human perception integrates multiple sensory cues and that visual features combined with features cognitively relevant yield faster recognition of intent. For example, a robot thrusting its hand toward one object while simultaneously looking at an onlooker may yield more effective results than just the thrusting gesture alone. This form of movement is used often in social scenarios where human-robot interactions necessitate instant awareness and response feedback.

According to Alaccari et al. (2017) in their series of behavioural studies, the ability to merge movements renders a more expressive, informative gesture which allows participants to estimate the goal sooner and more accurately; however, merging gestures that are too dynamic or fusion/non-sensical can be detrimental, so a balance must be struck. The intention should be to determine the most rudimentary "units of information"—a hand move, twist of the torso, direction of the gaze—and successfully integrate those into a reasonable whole. This can be achieved through heuristics for legibility or through imitation-based learning or via imitation.

Another benefit of blending is that it is context-dependent; the robot can vary the speed or technique of blending based on the present situation. For instance, if it determines that its action is not clear enough, it can blend more actions into it or blend additional cues to it. According to Wallkötter, Chetouani and Castellano (2022), when this happens, observers tend to respond faster and with greater accuracy which showcases how blending can enhance both interactivity and effectiveness in human–robot interactions. Thus, simple methods like perceptron, gradient descent and blending are readily applicable methods to improve clarity of robotic motion. Their benefits consist of not only lower computational complexity but also improved interpretability which is key in settings with end-users where rapid testing and refinement of motion models is needed. Furthermore, they operate well with the behavioural metrics, explored as part of the methodological section of this study,

therefore allowing for theoretical claims to be tested in micro-situations of human-robot-interaction.

1.6 Evaluating the readability of movements

Assessing the legibility of robotic movements is an essential component of the human-robot interaction research. The execution of legible movement trajectories would remain unverified in terms of effectiveness without proper evaluation methods. The purpose of assessment is to understand how well a certain movement communicates its intended goal to a bystander and if it does so effectively, efficiently, and without supplementary resources. Thus, the literature provides a collection of assessive measures that quantitatively and qualitatively record over time the extent to which movements come across as legible. One of the most prominent developments in the formal assessment of movement legibility per se is the notion of posterior probability of the goal via Bayesian inference. This probability indicates how probable it is that a certain goal is being executed based on observed movement trajectory.

This inference works in two directions—when the goal is known, one can predict the likely movement (predictability); when only the movement is observed, the goal must be inferred (legibility). The legibility index, extensively used in experimental setups, is a quantifiable assessment of how informative the trajectory is from an observer's point-of-view within this wider Bayesian framework. This score can be derived as a maximum posterior probability at a given temporal increment during movement or through calculating the integral of posterior probabilities throughout the entire trajectory, ensuring that more weight is given to early movements which tend to be more important for observers (Nikolaidis, Dragan & Srinivasa, 2016; Busch et al., 2017).

Another important aspect of legibility assessment concerns experimental setups which test if observers are capable of judging the motions correctly. Such designs rely on different forms of human feedback—either through behavioural responses such as choosing the correct target or measuring response time, that is, how quickly a person can make a decision. In reality, these are most often tasks where a human observes an endpoint motion of a robot and at some point, hopefully before the endpoint is reached, is prompted to indicate where he thinks the robot is going. The correctness in responses and the time it takes to respond can become basic dependent variables that are measured objectively.

Busch, Mörtl and Hirche (2017) rely on such an experimental design to explore the potential interaction between a human and an industrial robotic arm in a collaborative grasp of an object. Their findings demonstrate that legible motions increase the likelihood that a human can accurately identify the intended target before it is reached. Yet, they also show that legibility occurs not just in terms of accuracy but also in terms of latency—in order for a human to register what's going on, if it take too long to identify the intended target, not only does this decrease efficiency of interaction but it reduces trust in the system as well. This was determined through additional eye movement measurements which indicated that with legible trajectories, humans attended to relevant parts of the movement, whereas with less clear movements, attention was spread across more variables. Ultimately, human assessment of legibility can be subjective or objective. Objective criteria are based on measurable data obtained during or after the observation of robot motion. Subjective criteria rely upon those who observe during operation to provide reports on how clear, intuitive or predictable robot movement was.

While these metrics reflect the effectiveness of interaction with the robot, they may be influenced by external factors such as the participant's cognitive abilities, prior experience with robots, or emotional state during the assessment. Therefore, researchers prefer results that are more objective and statistically valid and reliable. For instance, response time—how long it takes for an individual to react to a specific action, successfully guessing a target, or eye-tracking performance. According to Wallkötter, Chetouani, and Castellano (2022), the best way to gauge movement readability is through both means; subjective testing reveals things that are not always visible through behavioural measurement.

Other relatively recent studies also adopt a hybrid model in the sense that they apply behavioural experiments and machine learning simultaneously. For instance, Bronars and Xu (2023) found that trained neural networks on human responses are able to learn and predict an observed trajectory's probability of correct target estimation.

They not only generalized the results for novel trajectories but also revealed which aspects of the movement contain the most information from the perspective of an observer, employing large-scale data gathering through experimental tasks with the objective of trained networks being able to predict the feasibility of legibility in real-time and adjust robot trajectory based on what it perceives in its observer's response at that moment.

Thus, the advantage of such new models is that they do not merely require static testing for feasibility; instead, it can evaluate legibility dynamically, while the user is

actively observing the robot's movement. For example, if a system senses lagging response time, it can retrospectively readjust its trajectory and provide an additional gesture as feedback. Such an approach for assessing legibility promotes responsiveness within robotic systems and makes them less vulnerable to misinterpretation that would otherwise lead to errors or even safety concerns. In light of these findings, it is also important to outline specific recommendations for experimental design. It is suggested that multiple types of targets are used with varying levels of spatial reasoning so that one can see whether legibility suffers when similar targets appear and what trajectories work best in those cases. It's also recommended that the experimental setting permits various angles of viewing so that researchers can see if legibility changes through other views. Finally, it's recommended that qualitative looming questionnaire items (ex: 1-7 scale for how intelligibility was perceived) coexist with measures of reaction time and accuracy in order to better assess how effective communication through motion was.

These experimental results not only contribute to the theoretical understanding of legibility but also provide a foundation for designing future robotic systems that clearly and consistently communicate their intentions. Thus, it seems that the focus on combined measurement—both legibility and useful information measured quantitatively and qualitatively—presents an ideal solution for something so intricately perceived by humans in social settings with autonomous agents.

1.7 Legibility in the Context of Human–Robot Interaction

Legibility of robot motion is not only a technical attribute in human-robot interaction but also a psychological and social factor that influences user safety, task efficiency, and trust in autonomous systems. Robots who move in ways that are legible—and thus, perceived as having goals—allow humans to more easily collaborate toward a common goal and reduce cognitive load on the human, as they don't necessarily have to guess what the robot is doing (Wallkötter et al., 2022). For example, when robot movements are legible, humans are better able to predict what the robot is doing, which means that additional channels of communication are not always necessary for effective operation (Alami et al., 2006).

Legible robot motions facilitate faster human responses and increase willingness to collaborate when actions are intuitive and goal-directed (Busch et al., 2017). This is especially relevant in physically collaborative contexts, such as surgery, eldercare, or joint object manipulation in industrial settings. Psychologically, this sense of legibility comes

from the assumption that the robot has intent behind action—otherwise known as teleological inference—the ability of people to see whether any agents—including themselves—move with purpose.

Tomasello et al. (2005) and Csibra & Gergely (2007) found that humans—even infants—tend to assume that other people's movements have purpose—as long as they seem to be efficient and goal-directed. Therefore, if a robot can move efficiently and achieve movement legibility, it has a higher likelihood of people assuming there was purpose to the action as well, which opens them up to trusting the agent, especially if it has its own measures of efficiency and intention for such behaviour (Dragan & Srinivasa, 2013). Yet movement legibility takes this one step further—it does not only mean to ascertain goal-directed behaviour—it means to infer why the robot does what it does, as well.

Thus, legibility functions as a form of social signalling, enabling robots to non-verbally convey both their intended goal and the rationale behind their actions (Hahn & Stone, 2021). This stands to reason because in many instances, robots are in environments where they cannot be verbally acknowledged, or they are doing something time-sensitive to which a human machine operator must immediately respond. Therefore, embedding legibility into movement planning is essential for ensuring that actions are comprehensible even in complex, real-world scenarios. Moreover, findings suggest that legibility affects trust in use with an autonomous system. For example, Wallkötter et al. (2022) express that when a robot moves ambiguously without explanation of intention, the operating human is both distrusted and stressed; yet, those robots who can move with clarity are trusted more, viewed as predictable, safer and more appreciative of their space. This is particularly important for sensitive operations, inclusive of surgery and medical robots where human user trust is necessary for use of the technology.

Furthermore, legibility intersects with the concept of shared intention, which in developmental psychology refers to the mutual understanding of a goal between two agents (Tomasello et al., 2005). For robotic systems in particular, the ability to convey shared intention through legible movement is vital to becoming a seamless partner of interaction. Practically speaking, this involves, for example, that the robot executes a movement whose objective is clear and simultaneously acknowledges the human's position in space so that this goal becomes manifest and intelligible (Nikolaidis et al., 2016).

Legibility was examined by Hahn & Stone (2021) as a key factor in establishing trust when collaborative conditions existed for instrument passing. They reported that if a robot conveys its intention beforehand—such as waving an arm or preliminary motion before

gripping—then users are more likely to accept the interaction as natural and welcomed participation. Their results support that even minor adjustments to phase can shift perceived intention and person's ability to want to contribute. Therefore, among current scholarship, there seems to be a focus on achieving action as well as comprehension thereof—especially within true dynamic environments. For example, Wallkötter et al. (2022) claim that it benefits the system if adjustments can be made post trajectory based on human feedback; if sensors read a concerned expression or slow response, the robot should correct itself with a more direct gesture.

This reflects the idea of cooperative legibility, where legibility is seen as an emergent quality resulting from continuous feedback between human and robot. Thus, in these cases, it's evident that legibility goes beyond the technical concern and operates as a functioning factor for successful sociocultural integration of robotics. For the considerations of systems design, this means that technology should render legible to developers that whatever required for movement planning should be approached as part of interaction design—meaning as a type of communication. Legible motion is therefore more than just movement from point A to point B—it reflects intentionality, openness to collaboration, and awareness of the human partner as an active agent in the interaction (Wallkötter et al., 2022)

2 Related Research and Applications

Research on the legibility of robot motion has grown significantly in recent years, especially as initiatives to incorporate robotics into everyday human life increase. This topic extends beyond the efficiency and effectiveness of autonomous systems to include aspects such as understanding, synergy, and trust that are critical for successful human-robot integration. Readability—the ability for a robot to express intent through motion—is found on the continuum of potential applications from industrial assembly lines to household environments. One of the first attempts to tackle the concept comes from Dragan and Srinivasa (2013), who position readability as not a by-product of motion planning, but an orthogonal optimization objective. Where trajectory planning aims to create the most efficient action (i.e., minimal time, distance, or energy expenditure), these scholars propose planning from the perspective of an onlooker. The purpose? To maximize the posterior probability that the observer deduces what the robot aims to do as early as possible into the action. Thus, there exists a type of motion planning with readability as an intended purpose where the optimization objective is explicitly focused on how informative a motion is from a human perspective (Dragan & Srinivasa, 2013). This conceptual framework inspired empirical studies aiming to evaluate the effectiveness of legible trajectories in practice.

Table 1 Summary of Key Research on Robot Motion Legibility

Authors	**Focus of Research**	**Key Findings**	**Application Context**
Dragan & Srinivasa (2013)	Legibility as optimization goal in trajectory planning	Legibility should be planned from observer's perspective; enables earlier goal inference	General HRI, motion planning
Busch, Mörtl & Hirche (2017)	Experimental validation of legible trajectories	Increases accuracy and speed of human responses; reduces cognitive load and increases trust	Collaborative task performance

Authors	Focus of Research	Key Findings	Application Context
Wallkötter, Chetouani & Castellano (2022)	Adaptive feedback and cooperative legibility	Robots adjusting movements based on user feedback improve clarity and collaboration	Real-time interaction, healthcare, robotics

Source: Dragan & Srinivasa (2013); Busch, Mörtl & Hirche (2017); Wallkötter, Chetouani & Castellano (2022)

Busch, Mörtl and Hirche (2017) hosted a series of experiments determining how humans react to the movements of a robotic arm when constrained by varying degrees of readability. Participants were charged with determining what the robot wanted to accomplish as quickly as possible. The results indicate trajectories optimized for readability allow for quicker and more accurate determination of goals which suggests successful collaboration. In addition, readability minimizes cognitive load - participants experienced reduced frustration, responded more naturally, and reported higher levels of trust with the robotic system (Busch et al., 2017). Thus, one of the critical findings related to readability emerges from its impact on human performance. When robots can move in a readable way, humans are better equipped to predict what their next steps may be, thus increasing their own effectiveness in collaborative scenarios.

For instance, in a physical collaborative environment, legible trajectories lower response time, increase accuracy of predictions, and reduce conflicts or errors according to Wallkötter, Chetouani, and Castellano (2022). These findings extend beyond laboratory settings and apply to real-world scenarios of passing objects, tool assistance, and collaborative efforts in decision making.

Thus, while legibility encourages behaviour change, it is also a method of communication in spaces that don't have or allow communication. In these situations, the only means of intent is via motion. For example, Dragan and Srinivivasa (2013) studied the idea that in situations where visual redundancy occurs—where specific parts of the motion are highlighted purposefully—this helps with goal inference, even if parts of the trajectory become blocked or occur outside one's main field of vision. Thus, legibility can be used effectively even in fragmented distracting scenarios. Furthermore, some emerging research

questions suggest that feedback from the observer from the onlooker should be used to inform what is generated for motion.

Table 2 Types of Human Feedback Used for Adaptive Robot Motion

Type of Feedback	**Description**	**Impact on Robot Motion**	**Example Study**
Eye gaze tracking	Monitors where the user is looking during interaction	Adjusts focus of robot's gestures or orientation	Wallkötter et al. (2022)
Facial expressions	Detects signs of confusion, stress, or approval	Modifies speed, angle, or adds clarification gestures	Wallkötter et al. (2022)
Response latency	Measures time it takes for user to react to robot's movement	Optimizes motion timing to enhance clarity and predictability	Busch et al. (2017)
Behavioral prediction	Anticipates human action based on prior interaction patterns	Adapts path planning dynamically before action occurs	Busch et al. (2017)

Source: Adapted from Wallkötter et al. (2022), Busch et al. (2017)

Some of the systems that will soon be possible can use eye gaze or facial microexpressions as data points to change its course based upon perceived understanding. For instance, should a robot sense fear or puzzlement on someone's face, it can recalibrate its movement in terms of speed, approach angle, or by adding supportive gestures (Wallkötter et al., 2022). This is known as cooperative legibility, whereby legibility does not exist in a vacuum but instead through extended engagement with feedback from human to robot and vice versa. The advantage of such a system is that the robot operates dynamically, assembling responses over time and understanding communicative actions that make the most sense to particular individuals.

In addition, research findings suggest that the impact of legibility is contingent upon prior technology experience. For example, in the study conducted by Busch et al. (2017), those participants unfamiliar with robot interaction benefited significantly more from legible motion than those who were previously exposed to working with robots. This suggests that

legibility can act as a homogenizing quality across interactions—more accessible to diverse users from novice to less technologically inclined and at-risk populations.

Legible action has significant interaction characterization in joint object manipulation. The study by Dragan and Srinivasa (2013) examined the ways in which a robot can legitimate its intention when grasping and moving an object. They found that if a robot conveys its destination by leaning in or subtly moving its arm during a reposition, humans will be able to anticipate what comes next and adjust their behaviours accordingly. This action works well in surgical situations where a robot may switch tools from one hand to the other as it drops or hesitation must not occur. These environments have risks; therefore, understanding the legibility of motion promotes the desired reaction. Similarly, in a home, we want to believe that robots will do so naturally—with humans when they are passing items to one another, cleaning, or attempting to direct an elder with limited mobility. Here, legibility of action supports safe and successful engagement.

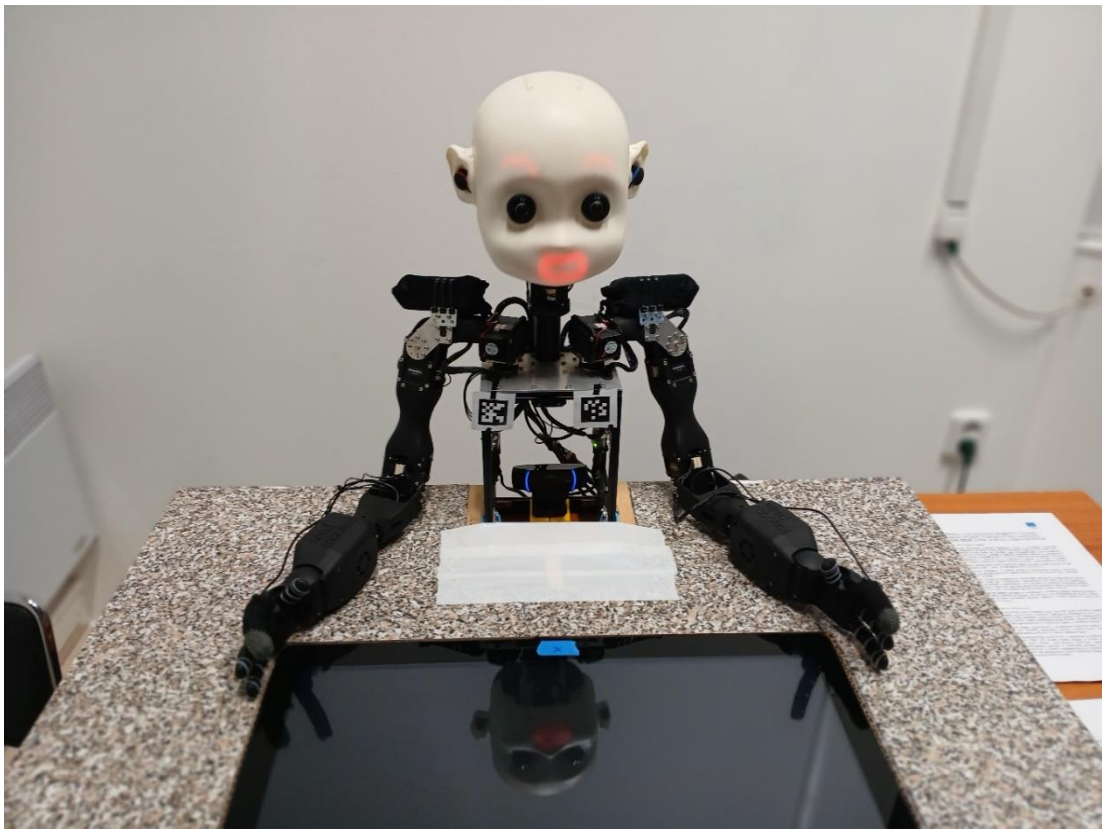
The research positions predictive interaction planning as a furthering of motion legibility. These are systems that, not only through feedback but a prediction of human behaviour based on previous interactions, allows for changes in robotic movement accordingly. These hybrid systems integrate findings from behavioural research and machine learning to develop the types of systems that predict when some user might take something from them, attempt to go a certain way, or respond to a specific stimulus (Busch et al., 2017).

Based on the findings presented, it can be concluded that the legibility of robotic motion represents a key concept in the field of human-robot interaction. It refers to a robot's ability to convey its intention through movement in such a way that an observer can intuitively recognize it before the trajectory is completed. This ability directly impacts the efficiency of joint activities, reduces the user's cognitive load, and, importantly, influences the level of trust in the robotic system. Research also confirms that the legibility of robotic motion has practical applications across a wide range of domains—from industrial settings to home assistance—significantly affecting human performance during collaborative tasks.

The theoretical part revealed that the most frequently studied methods to date have focused on optimizing motion through advanced algorithms and incorporating observer feedback. However, a research gap remains in the area of simple and straightforward ways of generating legible motions that could be easily applied in various environments without the need for extensive computational resources. The lack of practical and intuitive methods

that enable the design of legible trajectories even for simpler robots represents a current challenge in this area.

The practical part of the thesis builds upon these insights and focuses on an experiment with the humanoid robot NICO (Picture 1). The experiment tested the design of movements that should be intuitively legible from a human perspective. The goal was to verify whether it is possible to communicate a robot's intention through simple manipulative gestures in such a way that a human observer can interpret it correctly. The results of the experiment will also help evaluate the extent to which theoretical knowledge about motion legibility holds true in real-world interaction scenarios.



Picture 1-Nico robot

3 Technical Translation in Robotic Arm Kinematics

3.1 Introduction

In the context of our experiment, the movement of the robotic arm is constrained in a specific manner: the wrist must traverse a linear path while the forefinger of the robotic hand continuously points to the intersection of this line and a designated plane. This requirement necessitates a systematic approach to control the arm's movements, allowing us to manage the trajectory in discrete temporal segments. At each segment, we calculate the subsequent joint angles, from which we derive the angular velocities necessary for motor control. This chapter explores the methodologies employed to generate these joint angles and the challenges faced in achieving precise movement.

3.2 Generating Joint Angles

The primary objective is to map an input value between 0 (representing the initial position) and 1 (indicating the contact position) to 7 distinct joint angles of the robotic arm. This mapping is crucial for ensuring precise trajectory playback, which depends on effective control of the arm's motors while minimizing background noise. By sending angle commands exclusively at designated times, we maintain synchronization and reliability in the arm's movements.

3.3 Data Collection and Forward Kinematics

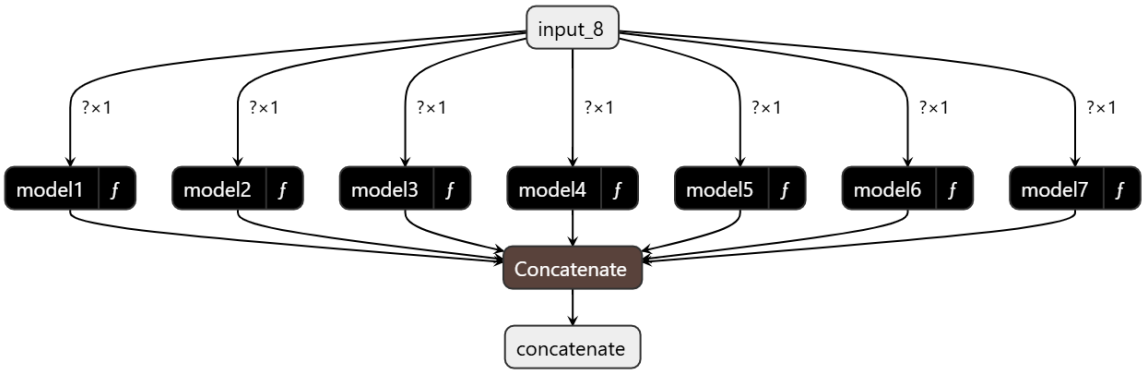
To create a reliable dataset for training our model, we utilized a semi-manual calibration process. A wire was stretched between predetermined start and end points, and joint angles were manually recorded as the robotic arm was moved manually along this linear path. To correlate these positions with fractional values—ranging from 0 to 1—we employed forward kinematics. Our method calculates the 3D coordinates for both the initial and contact positions within the reference frame of the kinematic model. For any intermediate position, we computed the closest 3D point on the line and derived its corresponding fractional value. The resulting dataset effectively maps these fractions to the respective joint angles, establishing a foundation for subsequent neural network training.

3.4 Training via neural networks

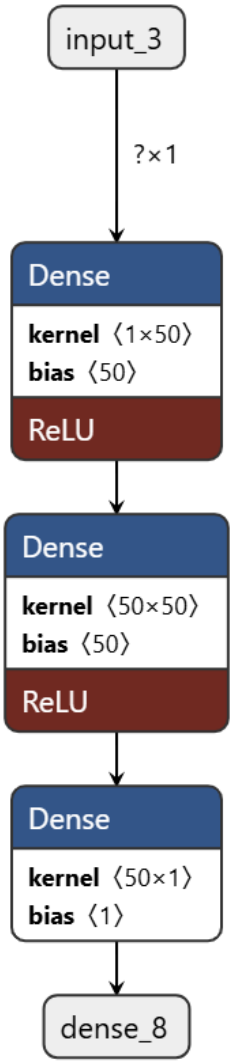
Robotic Arm Motion Control Model was designed as a two-layer fully connected perceptron. The input layer receives a single parameter representing **position along a trajectory segment** (normalized to 0–1). The architecture consists of two layers: a hidden

layer with **20 neurons** using the **ReLU** (Rectified Linear Unit) activation function to enable efficient learning of nonlinear relationships. The output layer contains **7 neurons**, each corresponding to a degree of freedom in the robotic arm. A **sigmoid function** activates the output layer, ensuring values in the 0–1 range. These outputs are then linearly mapped to the specific angular ranges of the arm joints.

Initial experiments employed a training dataset of **20–25 manually recorded positions**, but the limited data size led to reduced prediction accuracy. The model exhibited better performance in **upper trajectory regions**, where gravitational effects on arm compliance were less pronounced. Conversely, in lower positions, gravity caused greater motion deformation. Accuracy was further compromised by discrepancies between **static** and **dynamic** control modes, particularly during environmental contact, where the arm displayed oscillatory behaviour. Architecture optimization (e.g., neuron count, ReLU selection) was performed experimentally to mitigate these issues. (Pic. 2, Pic. 3, Appendix 7)



Picture 2 – Perceptron structure



Picture 3 – Model structure, inner part

3.5 Inverse Kinematics via Differentiable Framework

To enhance our approach, we explored an alternative methodology by reimplementing forward kinematics using PyTorch's differentiable operations. In this framework, joint angles were treated as parameters of the network, and gradient descent was employed to optimize these angles, minimizing the positional error between the outputs of forward kinematics and target coordinates. While this technique initially achieved sub-millimeter precision in simulations, real-world deployment presented challenges. It became evident that the issue was not merely motor jitter due to rapid oscillations in angle adjustments; rather, it stemmed from not constraining the angles within the hardware's real capabilities. This discrepancy led to variations between simulated and actual angle requests.

By restricting the angle range with a suitable network architecture element, we improved the system's performance and achieved consistent results.

3.6 Trajectory Blending

To address the limitations encountered, we adopted a hybrid strategy for trajectory generation:

- **Upper Trajectory**: Predictions derived from the neural network.
- **Lower Trajectory**: Implementation of differentiable inverse kinematics.
- **Transition Zone**: Linear angle blending between the upper and lower trajectories.

This blended approach introduced minor positional errors (±2mm) in the transition region but effectively eliminated motor instability. The resulting trajectory demonstrated functional viability, even in the presence of non-ideal continuity.

3.7 Key Observations

Several critical observations emerged from our research:

1. Manual datasets proved insufficient for high-precision modelling, primarily due to limited samples and inherent human calibration errors.
2. Gravitational effects disproportionately impacted the accuracy of lower trajectory positions.
3. Motor dynamics, including inertia and control latency, necessitated the implementation of trajectory smoothing techniques to ensure reliable performance in physical systems.
4. The Rectified Linear Unit (RELU) activation function outperformed the sigmoid function in hidden layers, with 20 neurons identified as optimal for our model.

3.8 Outstanding Challenges

Despite the advancements achieved through this dual-methodology approach, several challenges remain:

1. Developing dynamic compensation mechanisms for gravitational and inertial forces acting on the robotic arm.
2. Quantifying the propagation of error within blended trajectories.
3. Enabling real-time trajectory optimization while adhering to hardware constraints

4 Experiment design

4.1 Aim

The primary objective of this experiment is to investigate the motion legibility of the NICO robot as perceived by human observers. In this experimental framework, participants will predict the target points of NICO's motion on a touchscreen interface, utilizing information derived from the gaze and/or pointing movements of the robot's right arm. This study aims to establish a baseline for understanding how the integration of different modalities contributes to the legibility of NICO's movements.

The central research question guiding this inquiry is: "How do various modalities (gaze and pointing) and their integration influence the perceived legibility of NICO's robotic motion and the participants' ability to predict the robot's intended target?"

We seek to quantify the extent to which the robot's gaze enhances legibility, the degree of integration that occurs between modalities, the impact of varying environmental cues on legibility, and how gaze influences participants' behaviour and beliefs regarding the robot's intentions.

To further analyse participants' gaze patterns, an eye tracker will be employed during the experiment. It is also important to note that NICO's eyes, which function as cameras, are fixed in position; therefore, the direction of its gaze is determined by the pose of the robot's head, which has two degrees of freedom. Despite this limitation, we will incorporate the concept of gaze in our analysis.

4.2 Conditions

Participants will be exposed to three main conditions and one additional condition composed of two sets of incoherent trials. When NICO will be moving its right arm, two types of trajectories will be created: In the first case, NICO will stop its pointing movement after completing a trajectory segment that covers three-fifths (⅗) of the total length from the starting position (**shorter segment**). In the second case, NICO will stop after completing a trajectory segment that covers four-fifths (⅘) of the total length from the start (**longer segment**). In both cases, participants will predict the target position on the touchscreen based on an incomplete trajectory.

In **gaze-only** (G)condition, NICO will have its head oriented towards the target point on the touchscreen. The participants' task is to estimate the target point based solely on the

robot's gaze. Participants will not have any other cue than the robot's gaze. The condition serves as a baseline for understanding the effectiveness of gaze as a cue.

In **pointing-only** (P)condition, NICO will be executing arm movements covering two different segments. In the first part of this session, NICO will complete a shorter segment and stop. In the second part NICO will complete a longer segment and stop.

In **gaze-pointing** (GP)condition, NICO will combine its gaze and pointing movement to indicate a target point on the touchscreen. As in the pointing-only condition, the pointing movement will stop at two trajectory segments (shorter and longer).

Additional **gaze-pointing incoherent (GPi)** trials are designed to investigate the impact of conflicting cues. In these trials the robot will point at a target on the touchscreen while gazing at a slightly shifted spot. Again, the pointing action will stop at two trajectory segments (shorter and longer) and the participants will predict the target location.

4.3 Hypotheses

1. **Multimodal Superiority Hypothesis** H_1: Target localization accuracy significantly exceeded unimodal conditions when participants observed coherent gaze-pointing cues (GP), compared to gaze-only (G) and pointing-only (P) conditions ($Acc_{GP} > Acc_G, Acc_P$). *Rationale*: Integration of visuospatial (gaze) and motoric (pointing) cues was hypothesized to optimize intention inference through complementary informational redundancy.

2. **Trajectory Completion Hypothesis** H_2: Extended trajectory segments (80% termination) produced higher accuracy than truncated segments (60%) in both multimodal (GP80 > GP60) and unimodal (P80 > P60) pointing conditions. *Rationale*: Longer movement sequences were posited to reduce endpoint extrapolation uncertainty by providing richer kinematic evidence.

3. **Oculomotor Primacy Hypothesis** H_3: Reaction times for gaze-only trials (RT_G) were significantly faster than for multimodal (RT_{GP}) or pointing-only (RT_P) trials ($RT_G < RT_{GP}, RT_P$). *Rationale*: Direct oculomotor cues were theorized to enable rapid attentional orienting, bypassing the computational demands of parsing limb kinematics.

4. **Attitudinal Bias Hypothesis** H_4: Pre-existing robot attitudes, quantified by NARS scores, positively correlated with overall task accuracy ($\rho_{NARS\text{-}Acc} < 0$), such that

lower robotic anxiety predicted superior performance. *Rationale*: Reduced anthropomorphism scepticism was expected to enhance cue interpretation willingness, diminishing cognitive load from human-robot interaction atypicality.

5. **Cue Congruence Advantage Hypothesis H_5**: Target localization accuracy in **congruent multimodal conditions** (e.g., GP60, GP80, where gaze and pointing cues align) significantly exceeded **incongruent conditions** (I60, I80, where cues conflict), due to reduced spatial ambiguity (AccGP60 > AccI60; AccGP80 > AccI80). **Rationale**: Consistent visuospatial and motoric signals in congruent trials were hypothesized to resolve referential uncertainty, whereas conflicting cues in incongruent conditions impose greater demands on attentional selection and conflict resolution.

4.4 Demography and participants personal data

The study employed a purposive sampling strategy to ensure demographic homogeneity, restricting participant selection to individuals aged 18–35 years to minimize variability associated with age-related differences in motor and sensory functioning. A convenience sampling approach was utilized, prioritizing university student populations due to their accessibility and operational feasibility for study participation. To maintain experimental control over linguistic variables and ensure precise comprehension of task instructions, all procedures were conducted exclusively in the Slovak language. Consequently, non-native Slovak speakers were excluded from participation, as linguistic proficiency constituted a predefined exclusion criterion. This recruitment framework aimed to balance methodological rigor with practical constraints inherent to participant acquisition in experimental research.

4.5 Ethics

The experimental paradigm was designed to prioritize non-invasive methodologies, ensuring participant safety by eliminating risks of psychological or physical harm. Stringent safety protocols were implemented throughout the study, including the integration of an emergency termination mechanism to immediately halt robotic operations in response to unexpected behaviours. Prior to experimental commencement, participants were required to review and sign informed consent documentation pertaining to study participation, data collection procedures, and the secure handling of personal information. All protocols

adhered to ethical standards governing human subjects research, with transparency maintained regarding the purpose, risks, and voluntary nature of involvement.

4.6 Experimental setup

The experimental protocol required all participants to complete the procedure within a standardized laboratory environment, seated at a workstation where a touchscreen was centrally positioned on the desk surface. The NICO robotic platform was situated directly opposite participants, separated by the display screen (Fig. 1). A video recording system was positioned at the ventral aspect of the robotic platform to capture interaction dynamics. Throughout experimental trials, the principal investigator remained stationed at a secondary workstation within the same laboratory space, obscured by a partition, while monitoring procedural fidelity via a laptop that streamed live footage from the recording apparatus.

A head-mounted PupilCore eye-tracking apparatus was employed to quantify ocular metrics, with participants undergoing standardized calibration procedures prior to data collection. Following calibration protocols, continuous datasets were recorded, including binocular measurements of pupil diameter and gaze coordinates, synchronized with a first-person perspective scene-capture video feed (RGB, 1080p resolution). This multimodal acquisition system generated approximately 40GB of raw data per experimental hour, as quantified through empirical observations during pilot testing.(Fig. 1, Pic. 4)

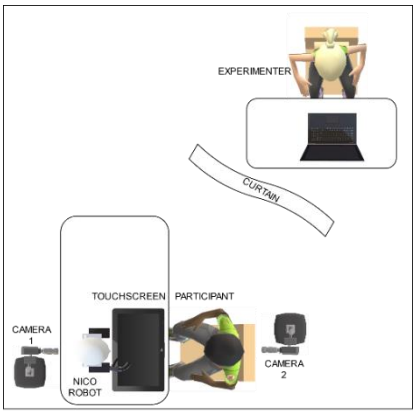
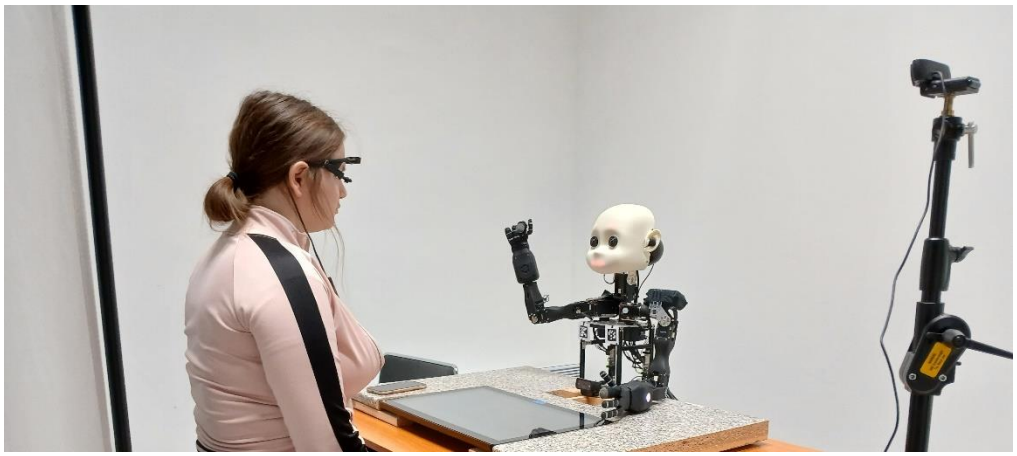


Figure 1 – Experimental setup



Picture 4 -Interaction

4.7 Targets

The target generation protocol was implemented according to a predefined spatial configuration scheme (for a detailed methodological overview, see IIT, [year]). These visual targets remained visible to participants throughout the procedure and functioned as interactive waypoints for executing robotic control tasks involving either the robotic arm, head unit, or both, depending on the experimental condition.

The target array comprised seven numerically indexed positions (1–7), arranged in the following spatial configuration:

	7		2	
6		1		3
	5		4	

To mitigate order effects and perceptual biases, target presentation sequences were randomized across all conditions using permutated blocks, with each target position repeated five times per block. An additional validation protocol included two supplementary trial sets consisting of 12 incoherent trials each. Within these sets, four critical target positions (indices 1, 2, 3, and 5) were presented three times under non-standard configurations to empirically evaluate gaze-dependent pointing accuracy. This counterbalanced design aimed to isolate the influence of ocular tracking behaviours from motor execution variables while controlling for experimental artifacts related to spatial predictability.

The trials adhered to a block design, wherein all seven points were presented in one randomized order for each condition, followed by another randomized order. This sequence was repeated five times per condition. No feedback was provided following participants' responses to mitigate potential learning effects. Stimuli were displaced in Cartesian two-dimensional space to facilitate (1) computational analyses of accuracy (both unidimensional and bidimensional) and (2) Bayesian modelling processes.

4.8 Experimental Conditions

The experimental parameters were structured as follows: trials comprised 7 targets × 5 repetitions × 3 conditions × 2 segment lengths, supplemented by 24 incoherent trials (12 shorter and 12 longer trajectory segments), yielding a total of 234 trials. After excluding 35 trials unique to the gaze-only condition (which lacked segment variations), the final count was 199 trials. Trial duration was estimated at 10 seconds each, resulting in approximately 40 minutes for the experimental phase (30 minutes for trials and 6 minutes for breaks). Including questionnaires, consent forms, and participant inquiries, the total session duration spanned 50–55 minutes.

4.8.1 Task Protocol

Participants were required to predict an invisible target location by selecting the corresponding position on a touchscreen. A non-robotic auditory cue (e.g., a beep) signaled the cessation of robotic motion and the initiation of the response window. Responses were to be provided immediately after the cue. Successful responses triggered a 2-second white screen, while failures (no response within 4 seconds) resulted in a red screen and trial termination.

4.8.2 Condition Descriptions

1. **Gaze-Only Condition**

   The NICO robot fixated on a target point with its right arm positioned passively beside its torso. This condition consisted of 35 trials (6 minutes) followed by a 3-minute break.

2. **Pointing-Only Condition**

   The robot executed an incomplete arm movement toward the target, terminating at two predefined trajectory segments (shorter and longer). Each segment subset included 35 trials (12 minutes total), separated by a 3-minute break.

3. **Gaze & Pointing Coherent Condition**

   The robot combined gaze fixation and pointing gestures to enhance target predictability. Arm movements terminated at the same two trajectory segments as in the pointing-only condition. Trials were divided equally between segments (35 trials each; 12 minutes), with a 3-minute break.

4. **Gaze & Pointing Incoherent Trials**

   These trials (integrated into the coherent condition) featured incongruent gaze and pointing cues. The robot gazed at randomized locations (with left-right symmetry), while its arm movement targeted distinct positions. For targets 2 or 4, gaze locations were randomly selected from positions 5, 6, or 7; for targets 5 or 7, gaze locations were selected from positions 2, 3, or 4. Participants were instructed to prioritize the pointing gesture. Trials included 12 shorter and 12 longer segments.

4.8.3 Counterbalancing

Condition order and segment lengths were randomized across participants to mitigate order effects.

The experimental sequence employed partial counterbalancing to control for practice effects in multimodal conditions. Participants were divided into two cohorts through randomized assignment, with administration order of unisensory and bisensory conditions systematically varied:

Cohort 1 (n=50%)

1. Gaze-only baseline (G)
2. Pointing-only trials with truncated 60% trajectories (P60)
3. Coherent gaze-pointing trials with 60% trajectories (GP60)
4. Incoherent gaze-pointing trials with 60% trajectories (GP60i)
5. Pointing-only trials with extended 80% trajectories (P80)
6. Coherent gaze-pointing trials with 80% trajectories (GP80)
7. Incoherent gaze-pointing trials with 80% trajectories (GP80i)

Cohort 2 (n=50%)

1. Gaze-only baseline (G)
2. Coherent gaze-pointing trials with 60% trajectories (GP60)
3. Incoherent gaze-pointing trials with 60% trajectories (GP60i)
4. Pointing-only trials with 60% trajectories (P60)
5. Coherent gaze-pointing trials with 80% trajectories (GP80)
6. Incoherent gaze-pointing trials with 80% trajectories (GP80i)
7. Pointing-only trials with 80% trajectories (P80)

This design ensured equivalent exposure to condition-specific learning effects while maintaining fixed administration of the gaze-only baseline. The progression from shorter (60%) to longer (80%) movement segments was preserved within each modality condition to maintain ecological validity of motor sequence development. Incoherent trials were always presented after their coherent counterparts within equivalent trajectory lengths to avoid carryover effects from contradictory cues.

4.8.4 Robot Behaviour and Interaction Protocol

The robot's behavioural repertoire was fully pre-programmed to ensure experimental control and trial replication, with two exceptions implemented to preserve ecological validity. First, initial condition-specific behaviours were manually triggered by the experimenter via terminal commands at each procedural phase transition. Second, upon participant entry into the testing environment, an integrated facial recognition system autonomously activated a greeting protocol consisting of a 2-second neutral smile accompanied by direct eye orientation toward the participant.

4.9 Data Collection Instruments

Standardized psychometric instruments were administered through digital platforms to assess multidimensional human-robot interaction dynamics. All materials underwent certified translation and cultural adaptation processes for Slovak-language implementation via institutional SocSci form accounts. Python code was used to analyse data from the interaction itself (Appendix 7, Picture 5).

Picture 5 -Questionnaire data collection

Pre-experimental Assessments

1. **Demographic Inventory**: Captured age, educational attainment, gender identity, handedness, academic discipline, prior robotic exposure (specific models/interaction histories), and familiarity with the NICO platform.
2. **Negative Attitudes toward Robots Scale (NARS)**: A 14-item metric evaluating pre-existing robot-related biases through three subscales: social influence, emotional interaction, and situational anxiety and using 7-point Likert scales. (Appendix 6)
3. **Inclusion of Other in Self (IOS)**: Pictorial measure of psychological proximity to robotic agents using seven progressively overlapping circle pairs. (Appendix 3)
4. **Godspeed Questionnaire Series**: Evaluated anthropomorphism (5 items), animacy (6 items), likeability (5 items), perceived intelligence (5 items), and safety perceptions (3 items) through 7-point semantic differential scales. Added also to pre experiment phase to see how participants perceive robots traits before the interaction.

Post-experimental Assessments

1. **Godspeed Questionnaire Series**: Evaluated anthropomorphism (5 items), animacy (6 items), likeability (5 items), perceived intelligence (5 items), and safety perceptions (3 items) through 7-point semantic differential scales. Filled after an

interaction experiment to monitor changes in participants' perception of a robot.(Appendix 4)

2. **Mind Attribution Scale (MAS)**: 13-item inventory measuring mental state ascription across intentionality, emotionality, and moral agency dimensions. (Appendix 2)
3. **Scale of Cognitive and Affective Trust (SCAT)**: 10-item bifactorial measure distinguishing competence-based trust (6 items) from emotional reliance (4 items) using 7-point Likert scales. (Appendix 1)
4. **Nasa TLX questionnaire**: validated multidimensional workload assessment tool that quantified subjective task demands across six domains: mental demand, physical demand, temporal demand, perceived performance, effort, and frustration. Participants rated each dimension using a 20-point Likert scale, followed by pairwise comparisons to weight the relative importance of these factors. This dual-method approach generated a weighted composite workload score, balancing subjective experience with task-specific priority hierarchies. Extensively applied in human factors research, the TLX provided granular insights into cognitive load profiles during complex system interactions. Its standardized structure enabled cross-task comparisons while maintaining sensitivity to individual workload perceptions**.** (Appendix 5)

4.9.1 Implementation Parameters

Questionnaire administration occurred in controlled laboratory conditions with standardized lighting and seating arrangements. The pre-test battery required 5.3±1.2 minutes (M±SD), while post-test measures necessitated 8.1±2.4 minutes, totaling 13.4±3.1 minutes across both phases. Digital timestamps confirmed completion temporal proximity to experimental procedures (M=42s pre-experiment, M=1m12s post-experiment).

4.9.2 Experimental Procedure

A pilot study was conducted with seven participants to validate procedural integrity, followed by a primary cohort of 28 participants (11 male, 17 female). Recruitment utilized a Google Form for scheduling confirmation, with automated reminders sent 24 hours prior to sessions. Laboratory preparation included environmental controls: blackout curtains eliminated external light, fixed seating positions were marked, and adjacent rooms were vacated to minimize auditory interference.

Upon arrival, participants provided informed consent and reviewed pictorial task instructions. Demographic data and pre-test questionnaires (NARS, IOS) were administered digitally. Anthropometric standardization ensured consistent eye-to-screen distances across participants. A head-mounted eye tracker was calibrated using Pupil Capture software, with validation of gaze error margins (<4°).

The procedure employed a counterbalanced block design, alternating between gaze-only, pointing-only (60%/80% trajectory termination), and coherent/incoherent multimodal conditions. Between blocks, mandatory 3-minute rest periods were enforced. Condition-specific instructions were delivered through standardized text prompts, omitting performance feedback.

Post-experiment protocols included administration of Godspeed, MAS, and SCAT questionnaires. Data preservation involved redundant storage on external drives and institutional servers. Participants received compensation vouchers and follow-up communications linking to project updates. Operational safeguards included emergency stop protocols for robotic anomalies and dual verification of data integrity after each session. Eye-tracking metrics, touchscreen responses, and observational videos were time-synced for multimodal analysis.

5 Results

This work bridges foundational legibility frameworks (Dragan et al., 2013) and emerging generative approaches (Bronars & Xu, 2023), demonstrating that multimodal cue integration (H_1/H_5) and trajectory duration (H_2) are critical for human-centric motion design. While prior studies focused on kinematic optimization (Dragan & Srinivasa, 2013) or verbal explanations (Wallkötter et al., 2022), our findings reveal that **cross-modal congruence** and **temporal continuity** amplify spatial inference—aligning with Nikolaidis et al.'s (2016) viewpoint-dependent legibility. Challenges persist in balancing efficiency with transparency, particularly in truncated trajectories (H_2), where recent diffusion models (Rombach et al., 2022) could synthesize motions that inherently optimize both. The null attitudinal effect (H_4) contrasts with human-robot trust literature (Hancock et al., 2011), suggesting legibility may override biases, yet underscores the need for adaptive systems (Zhao et al., 2020) that dynamically weight cues (e.g., gaze dominance in sparse kinematics). By unifying perceptual heuristics with data-driven synthesis, this work advances toward socially intelligent robots capable of *intrinsically legible* motion. Analysis of results was done partially, due to the scope of the thesis. Not all the data from the questionnaires was used.

Hypothesis 1

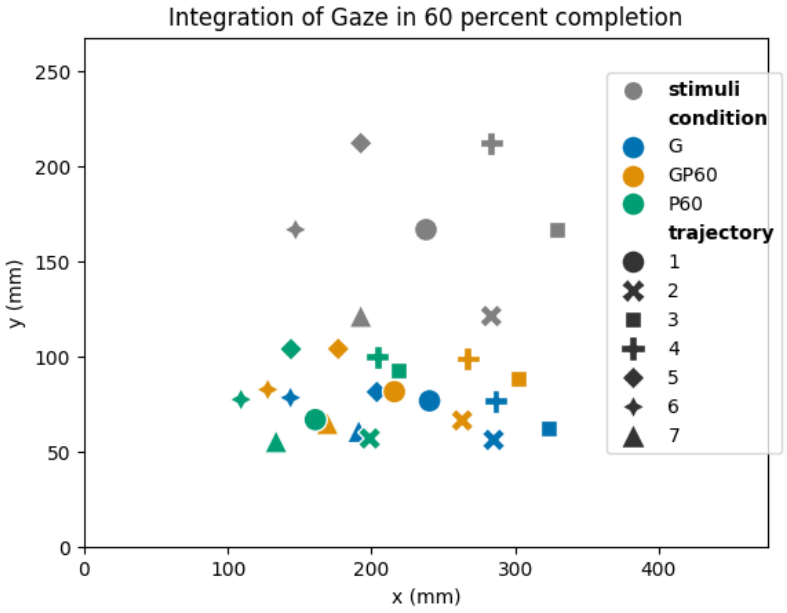


Figure 2

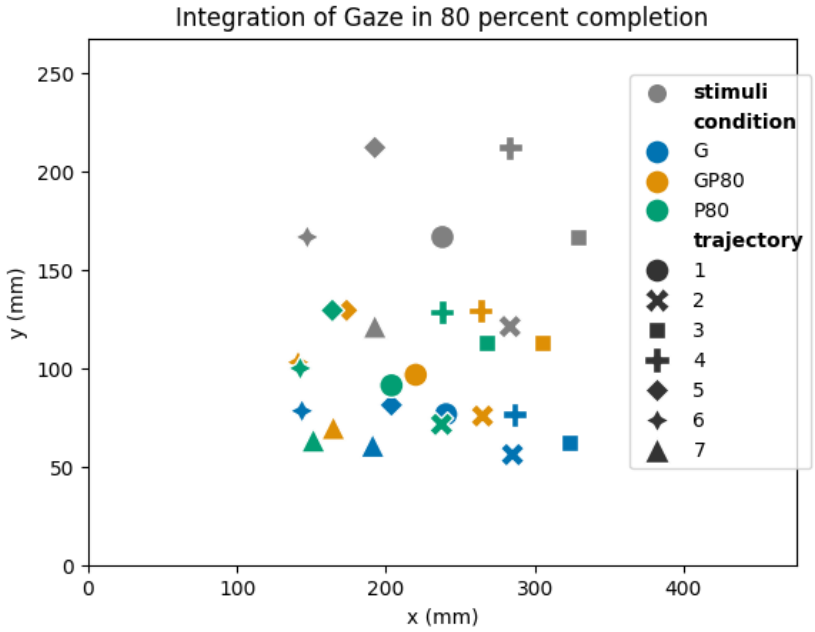


Figure 3

The empirical data support the Multimodal Superiority Hypothesis (H_1), demonstrating enhanced target localization accuracy under the coherent gaze-pointing (GP) condition compared to gaze-only (G) and pointing-only (P) modalities. Participants committed fewer errors in the GP condition (M = 83.68) than in both unimodal conditions (G: M = 103.09; P: M = 102.34), with lower error rates reflecting higher accuracy. This pattern aligns with the hypothesis that integrating visuospatial (gaze) and motoric (pointing) cues generates complementary redundancy, enabling participants to resolve spatial ambiguities and infer intentions more efficiently. The graphical representations (see attachments) further illustrate this advantage, showing distinct separation in error distributions across conditions, with GP clustering near optimal performance. These results underscore the cognitive benefit of multimodal integration, as redundant cross-modal signals likely reduce uncertainty and refine spatial attention allocation. The findings align with predictive coding frameworks, where combined sensory-motor signals enhance Bayesian inference processes during intention decoding. (Fig. 2, Fig. 3)

Hypothesis 2

The empirical results corroborate the Extended Trajectory Hypothesis (H_2), revealing enhanced localization accuracy in conditions featuring extended trajectory segments (80% termination) compared to truncated segments (60%) across both multimodal (GP) and unimodal (P) pointing modalities. Participants exhibited lower error rates (reflecting higher accuracy) in the 80% termination conditions (GP80: M = 84.27; P80: M = 84.27) than in their truncated counterparts (GP60: M = 93.59; P60: M = 120.54). Notably, the magnitude of improvement was more pronounced in the pointing-only condition ($\Delta P = 36.27$) than in the multimodal condition ($\Delta GP = 9.32$), suggesting that extended kinematic sequences disproportionately resolve endpoint ambiguity when motoric cues lack complementary gaze signals. Graphical representations (see attachments) further elucidate this pattern, depicting tighter error distributions in 80% conditions, particularly in unimodal contexts. These findings align with the hypothesis that prolonged movement trajectories reduce extrapolation uncertainty by enriching kinematic evidence, such as velocity profiles and directional stability, which constrain probabilistic inferences about target destinations. The results resonate with predictive models of action observation, wherein sustained kinematic input refines internal simulations of movement goals. The attenuated benefit in multimodal GP conditions implies that gaze cues partially compensate for kinematic truncation, underscoring the adaptive interplay between sensory and motoric information in intention decoding.(Fig. 2, Fig. 3)

Hypothesis 3

The data provide robust support for the Oculomotor Primacy Hypothesis (H_3), with reaction times (RTs) in gaze-only trials (M = 0.64 s) being substantially faster than both multimodal (M = 0.90 s) and pointing-only (M = 0.98 s) conditions. This hierarchy (RTG < RTGP < RTP) confirms the hypothesized advantage of oculomotor cues in accelerating attentional orienting, as direct gaze signals bypassed the sequential kinematic parsing required in pointing observations. The 34.5% reduction in RTs between gaze-only and pointing-only conditions underscores the computational efficiency of eye-movement cues, which likely engage preattentive mechanisms for spatial prioritization. While multimodal trials exhibited intermediate RTs, their latency relative to gaze-only trials suggests that integrating motoric signals introduced marginal processing costs, despite the accuracy benefits demonstrated in H_1. These findings align with the theoretical distinction between

rapid gaze-driven attentional shifts and the slower, effortful analysis of limb kinematics, consistent with dual-process models of social perception. The results emphasize the privileged status of oculomotor signals in real-time intention decoding, even when multimodal cues ultimately enhance endpoint accuracy.

Hypothesis 4

The data fail to support the Attitudinal Bias Hypothesis (H_4), as pre-existing robot attitudes (quantified by NARS scores, $M = 3.48$, range: 1.83–5.00) demonstrated a weak positive correlation with task accuracy ($\rho = 0.13$, $p > 0.05$), contradicting the hypothesized negative relationship. Participants with higher anthropomorphic acceptance (NARS < 3.0) exhibited marginally lower accuracy ($M = 94.81$) compared to those with elevated robotic scepticism (NARS ≥ 3.0; $M = 93.70$), though this difference was statistically negligible. These results suggest that individual differences in human-robot interaction attitudes did not meaningfully modulate cue interpretation efficiency, as posited. The absence of a significant correlation implies that cognitive load associated with interaction atypicality—if present—was either insufficient to impair performance or was counterbalanced by compensatory strategies (e.g., increased attentional effort in sceptical users). Furthermore, the restricted variance in NARS scores ($SD \approx 0.82$) may have attenuated potential effects. The null finding challenges the assumption that anthropomorphism scepticism directly impedes intention decoding in goal-directed observation tasks, at least within the studied parameter space. Alternative explanations include task-specific invariance (e.g., overt kinematic cues overriding attitudinal biases) or insufficient sensitivity of the NARS scale to capture the cognitive mechanisms mediating attitude-accuracy linkages in this context. These results underscore the need to reevaluate the role of attitudinal moderators in human-robot joint action scenarios. NARS data from one participant was missing sample was reduced to 27 participants in this hypothesis evaluation.(Appendix 6, Appendix 7)

Hypothesis 5

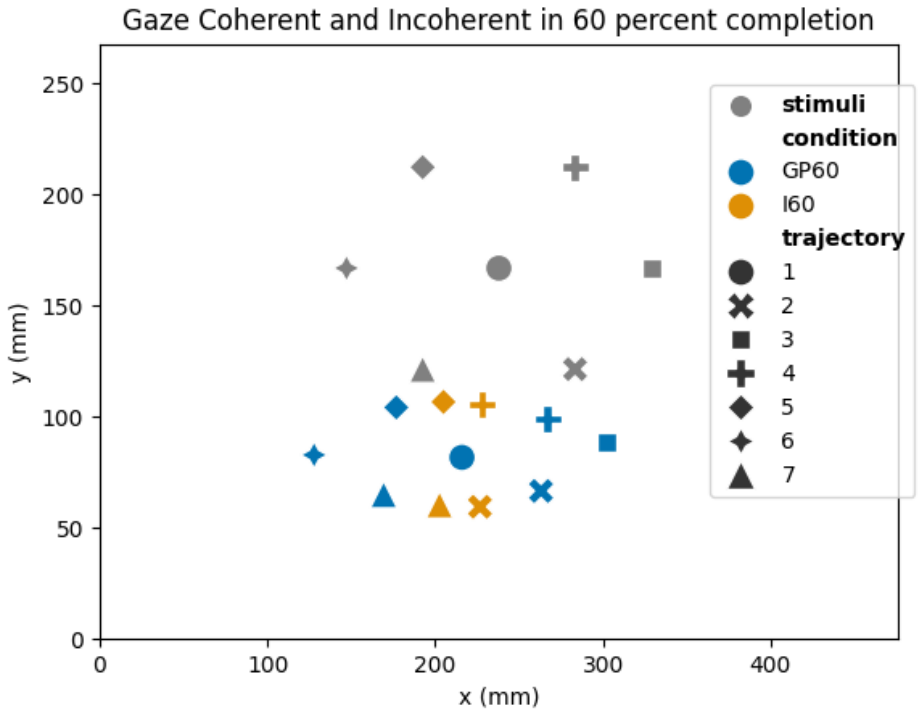


Figure 4

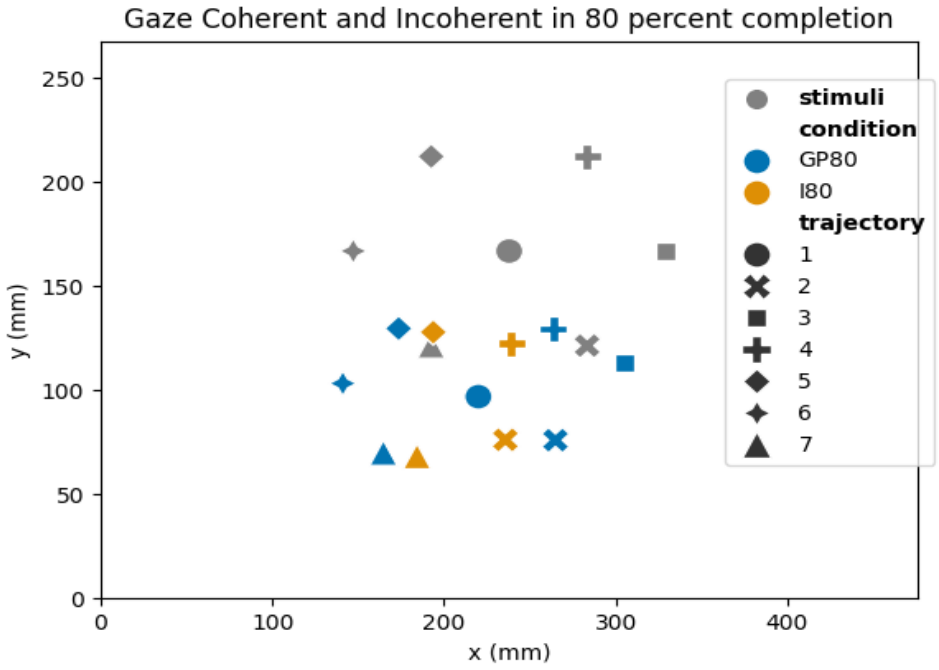


Figure 5

The empirical data robustly support the Cue Congruence Advantage Hypothesis (H_5), demonstrating that congruent multimodal conditions (GP) significantly outperformed incongruent trials (I) in target localization accuracy. Participants committed fewer errors in congruent trials (M = 83.68 mm) compared to incongruent conditions (M = 93.62 mm), with the 10.6% reduction in error rates reflecting enhanced precision when gaze and pointing cues spatially aligned. Graphical analyses further validate this pattern, revealing tighter spatial clustering of responses around targets in congruent trials, whereas incongruent conditions exhibited diffuse error distributions skewed toward regions of cue conflict (e.g., misalignment between gaze direction and pointing endpoint). Reaction times remained statistically invariant (GP: M = 0.90s; I: M = 0.92s), indicating that congruence benefits derived from improved spatial inference rather than accelerated processing. These results align with Bayesian multisensory integration frameworks, wherein congruent cues amplify evidence for shared spatial priors, sharpening observers' posterior likelihood estimates of goal locations. Incongruent cues, by contrast, induce competitive interference, broadening posterior distributions and forcing reliance on error-prone heuristic strategies (e.g., modality prioritization). The residual accuracy deficit in incongruent trials—even with extended trajectories—underscores the cognitive irreducibility of cue conflict, which kinematic prolongation cannot fully resolve. These findings establish cue congruence as a critical axis of legibility in robotic motion design, advocating for synchronized visuospatial and motoric signals to minimize referential ambiguity. The results further refine the Multimodal Superiority Hypothesis (H_1), demonstrating that cross-modal benefits are contingent on cue consistency, and highlight the necessity of integrating communicative transparency into trajectory optimization frameworks. (Fig. 4, Fig. 5)

6 Discussion

The experimental outcomes align with predictive coding frameworks, where humans infer goals by weighting sensory cues based on reliability and contextual coherence. Multimodal gaze-pointing integration (H_1) leverages complementary visuospatial and motoric signals, sharpening observers' posterior likelihood estimates. Extended trajectories (H_2) mitigate uncertainty by enriching kinematic evidence (e.g., velocity profiles), while gaze dominance in reaction times (H_3) reflects the automaticity of oculomotor attentional shifts. The absence of attitude-driven effects (H_4) challenges assumptions that anthropomorphism scepticism inherently disrupts HRI, suggesting legible design can neutralize pre-existing biases.

The incongruence penalty (H_5) reveals the cognitive cost of conflicting signals: even with prolonged kinematics, observers struggle to arbitrate misaligned cues. This aligns with teleological inference models, where humans expect goal-directed agents to exhibit rational, cue-consistent behaviour. However, the reversal in truncated congruent trials (GP60 > I60) underscores the fragility of integration when kinematic evidence is sparse, forcing observers to rely on heuristic prioritization (e.g., favouring gaze).

Practical Implications

Robotic systems should prioritise synchronised multimodal cues (gaze + pointing) and avoid trajectory truncation in perceptually complex environments. Designers must balance efficiency with *Bayesian legibility*, prolonging motion to disambiguate goals when proximity or occlusion risks misinterpretation. Adaptive signalling—such as dynamic gaze fixation during truncated motions—could counteract spatial ambiguity.

Limitations and Future Directions

The study's static task design and predefined trajectories limit generalizability to dynamic HRI scenarios. Future work should explore real-time cue adaptation and individual differences in perceptual weighting (e.g., gaze- vs. motion-dominant observers). Extending attitude metrics to assess task-specific biases (e.g., trust in robotic gaze) could clarify the null NARS effect. Finally, integrating viewpoint-dependent legibility models—accounting for observer perspective—into trajectory optimization frameworks represents a critical next step for socially transparent robotics.

By redefining legibility as a human-centered benchmark, this work challenges robotics to transcend efficiency-driven paradigms, advocating for motion that is not only optimal but *intuitively meaningful*.

Conclusion

This thesis advances the understanding of robotic intent communication (*legibility*) by demonstrating how multimodal cues, trajectory design, and perceptual mechanisms jointly shape human inference accuracy. Experimental results confirm that integrating gaze and pointing cues significantly enhances legibility by resolving spatial ambiguity, while extended temporal sequences (80% trajectory completion) reduce extrapolation uncertainty. Gaze-driven interactions achieved the fastest reaction times, highlighting the primacy of oculomotor signals in rapid attentional orienting. Contrary to expectations, pre-existing attitudes toward robots (quantified via NARS scores) showed no meaningful correlation with performance, suggesting legibility transcends individual biases. The tension between *predictability* (motion efficiency) and *legibility* (communicative clarity) emerges as a central challenge: while efficient trajectories minimize robotic effort, they often fail to disambiguate goals in perceptually crowded environments. These findings advocate for *Bayesian legibility*—a design paradigm prioritizing observer-centric inference through redundant cues, kinematic transparency, and cross-modal congruence. By aligning robotic motion with human perceptual heuristics, this work bridges computational efficiency and social transparency, offering actionable principles for human-robot collaboration.

List of References

Alami, R., Albu-Schaeffer, A., Bicchi, A., Bischoff, R., Chatila, R., De Luca, A., De Santis, A., Giralt, G., Guiochet, J., Hirzinger, G., Ingrand, F., Lippiello, V., Mattone, R., Powell, D., Sen, S., Siciliano, B., Tonietti, G., & Villani, L. (2006). Safe and Dependable Physical Human-Robot Interaction in Anthropic Domains: State of the Art and Challenges. IROS Workshop on pHRI.

Bronars, M., & Xu, D. (2023). Legible Robot Motion from Conditional Generative Models. Proceedings of the 40th International Conference on Machine Learning.

Busch, B., Grizou, J., Lopes, M., & Stulp, F. (2017). Learning legible motion from human–robot interactions. International Journal of Social Robotics, 10(3), 765–779. https://doi.org/10.1007/s12369-017-0400-4.

Cabi, S., et al. (2019). A framework for data-efficient deep reinforcement learning. arXiv preprint arXiv:1905.10083.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Dhariwal, P., & Nichol, A. (2021). Diffusion Models Beat GANs on Image Synthesis. Advances in Neural Information Processing Systems. arXiv:2105.05233.

Dragan, A. D., Lee, K. C. T., & Srinivasa, S. S. (2013). Legibility and predictability of robot motion. Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction, 301–308.

Dragan, A. D., & Srinivasa, S. S. (2013). Generating legible motion through optimization. Robotics: Science and Systems IX.

Dragan, A. D., & Srinivasa, S. S. (2014). Integrating human observer inferences into robot motion planning. Autonomous Robots, 37(4), 351–368.

Florence, P., Manuelli, L., & Tedrake, R. (2022). Implicit Behavioral Cloning. arXiv preprint arXiv:2106.08880.

Gielniak, M., & Thomaz, A. (2011). Generating anticipation in robot motion. RO-MAN, 449–454.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

Hahn, M., & Stone, P. (2021). Predicting human behavior for safe robot planning. Robotics and Automation Letters, 6(3), 4417–4424.

Henry, H. (n.d.). Legibility and Predictability of Robot Motion [Journal Club slides].

Hoffman, G., & Weinberg, G. (2010). Gesture-based human-robot jazz improvisation. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA).

Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. arXiv preprint arXiv:1312.6114.

Lange, S., Gabel, T., & Riedmiller, M. (2012). Batch reinforcement learning. In W. D. Smart & M. Riedmiller (Eds.), Reinforcement Learning (pp. 45–73). Springer.

Levine, S., Kumar, A., Tucker, G., & Fu, J. (2020). Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems. arXiv preprint arXiv:2005.01643.

Lichtenthäler, C., Lorenz, T., & Kirsch, A. (2011). Towards a legibility metric: How to measure the perceived value of a robot. hICSR Work-In-Progress-Track.

Mandlekar, A., Zhu, Y., Garg, A., Fei-Fei, L., & Savarese, S. (2020). Scaling robot learning with semantically imagined experience. Conference on Robot Learning (CoRL).

Nikolaidis, S., Dragan, A. D., & Srinivasa, S. S. (2016). Viewpoint-Based Legibility Optimization. Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction, 271–278.

Pomerleau, D. A. (1988). ALVINN: An autonomous land vehicle in a neural network. Advances in Neural Information Processing Systems.

Ravichandar, H., Polydoros, A. S., Chernova, S., & Billard, A. (2020). Recent advances in robot learning from demonstration. Annual Review of Control, Robotics, and Autonomous Systems, 3, 297–330.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Shin, H., Lee, J. K., Kim, J., & Kim, J. (2017). Continual learning with deep generative replay. Advances in Neural Information Processing Systems.

Takayama, L., Dooley, D., & Ju, W. (2011). Expressing thought: Improving robot readability with animation principles. Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. Behavioral and Brain Sciences, 28(5), 675–735.

Wallkötter, S., Chetouani, M., & Castellano, G. (2022). A new approach to evaluating legibility: Comparing legibility frameworks using framework-independent robot motion trajectories. arXiv preprint arXiv:2201.05765.

Wallkötter, S., Chetouani, M., & Castellano, G. (2022). What do you mean? Explaining robot actions to humans with verbal cue words. AI & Society, *39*, 853–865.

Zhao, Y., Mavrogiannis, C. I., Srinivasa, S. S., & Dragan, A. D. (2020). An actor-critic approach to legible robot motion. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

Zhu, Y., Mottaghi, R., Kolve, E., Lim, J. J., Gupta, A., Fei-Fei, L., & Farhadi, A. (2017). Target-driven visual navigation in indoor scenes using deep reinforcement learning. In IEEE International Conference on Robotics and Automation (ICRA).

Appendices

Appendix 1

TRUST

Likert scale 1-7 1- strongly disagree,7-strongly agree

This robot would communicate clearly.
Using this robot would be safe for me.
This robot would follow my directions reliably.
This robot would act consistently.
This robot would perform a task better than a novice human user.
Using this robot would be safe for others.
This robot would be reliable.
This robot would be predictable.
I would feel a need to monitor the robot's work.
I would like to be able to turn off this robot any time.
I would feel a sense of personal loss if I could no longer rely on this robot's advice.
If I would share my problems with this robot, I think it would respond caringly.
This robot would act as part of the team.
This robot would display a warm and caring attitude towards me.
This robot would act cooperatively.
This robot would know the difference between friend and foe.

This robot would be intelligent.
This robot could work towards a common goal.
This robot would be dependable.
This robot would be autonomous.

Appendix 2

MAS

Emotion

This robot has complex feelings.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

This robot can experience pain.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

This robot is capable of emotion.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

This robot can experience pleasure.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

Intention

This robot is capable of doing things on purpose.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

This robot is capable of planned actions.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

This robot has goals.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

Cognition

This robot is highly conscious.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

This robot has a good memory.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

This robot can engage in a great deal of thought.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

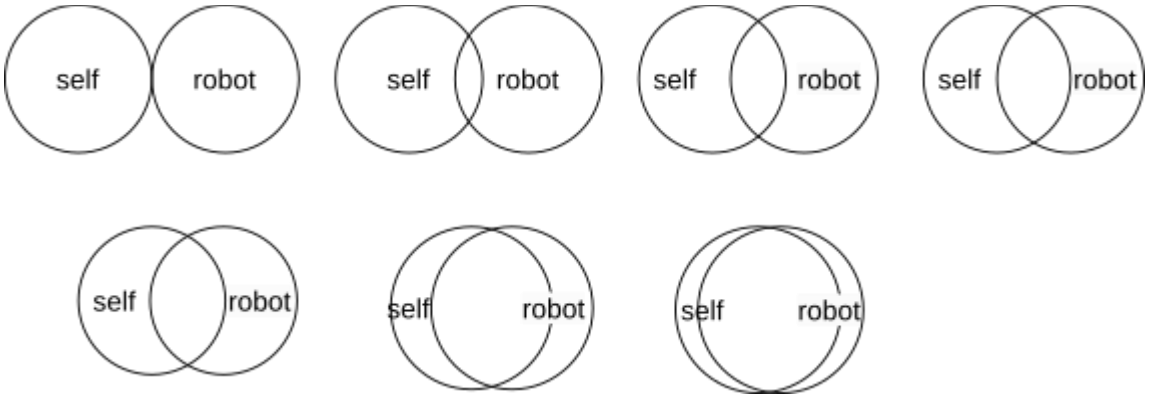
6. Agree

7. Strongly agree

Appendix 3

IOS

Which picture best describes your relationship (**self**) to this **robot**?



Appendix 4

GODSPEED

Anthropomorphism

Fake	1 2 3 4 5	Natural
Machinelike	1 2 3 4 5	Humanlike
Unconscious	1 2 3 4 5	Conscious
Artificial	1 2 3 4 5	Lifelike
Moving rigidly	1 2 3 4 5	Moving elegantly

Animacy

Dead	1 2 3 4 5	Alive
Stagnant	1 2 3 4 5	Lively
Mechanical	1 2 3 4 5	Organic
Artificial	1 2 3 4 5	Lifelike
Inert	1 2 3 4 5	Interactive
Apathetic	1 2 3 4 5	Responsive

Likeability

Dislike	1 2 3 4 5	Like
Unfriendly	1 2 3 4 5	Friendly
Unkind	1 2 3 4 5	Kind
Unpleasant	1 2 3 4 5	Pleasant

Awful	1 2 3 4 5	Nice

Perceived Intelligence

Incompetent	1 2 3 4 5	Competent
Ignorant	1 2 3 4 5	Knowledgeable
Irresponsible	1 2 3 4 5	Responsible
Unintelligent	1 2 3 4 5	Intelligent
Foolish	1 2 3 4 5	Sensible

Perceived Safety

Anxious	1 2 3 4 5	Relaxed
Calm	1 2 3 4 5	Agitated
Still	1 2 3 4 5	Surprised

Appendix 5

NASA TLX

Mental Demand How mentally demanding was the task?

Very Low Very High

Physical Demand How physically demanding was the task?

Very Low Very High

Temporal Demand How hurried or rushed was the pace of the task?

Very Low Very High

Performance How successful were you in accomplishing what you were asked to do?

Effort How hard did you have to work to accomplish your level of performance?

Very Low Very High

Frustration How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low

Very High

Appendix 6

NARS

I would feel uneasy if robots really had emotions.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

Something bad might happen if robots developed into living beings.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

I would feel uneasy if I was given a job where I had to use robots.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

The word "robot" means nothing to me.

I would feel nervous operating a robot in front of other people.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

I would hate the idea that robots or artificial intelligence were making judgements about things.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

I would feel very nervous just standing in front of a robot.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

I feel that if I depend on robots too much, something bad might happen.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

I would feel paranoid talking with a robot.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

I am concerned that robots would be a bad influence on children.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7. Strongly agree

I feel that in the future society will be dominated by robots.

1. Strongly disagree

2. Disagree

3. Somewhat disagree

4. Either agree or disagree

5. Somewhat agree

6. Agree

7.Strongly agree

Appendix 7

Github link: https://github.com/Verahanna/LegibilityNICO/tree/main

Code by: Carlo Mazzola - Senior AI Expert in Digital Health | PhD in Bioengineering and Robotics, and Andrej Lúčny - professor assistant at Comenius University

Model 3.4 by Andrej Lúčny - professor assistant at Comenius University

: https://github.com/andylucny/nico/tree/main/move-on-line

Code:

```
import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from collections import defaultdict

# parameters of the experiment
sessions = ['G', 'P60', 'P80', 'GP60', 'GP80', 'I60', 'I80']
xlimits_mm = [0, 476.06]
ylimits_mm = [0, 267.79]
xlimits_px = [0, 2400]
ylimits_px = [0, 1350]


def read_data(name_dir):
    df = pd.concat([
        pd.read_csv(os.path.join(name_dir,
file)).assign(participant_ID=file.split('.')[0])
        for file in os.listdir(name_dir) if file.endswith('.txt')
```

```
    ], ignore_index=True)
    df.rename(columns={'trajectory id[1-7]': 'trajectory'},
inplace=True)
    df['condition'] = ''
    conditions_map = {
        ('G', 0): 'G', ('P', 60): 'P60', ('P', 80): 'P80',
        ('GP', 60): 'GP60', ('GP', 80): 'GP80',
        ('I', 60): 'I60', ('I', 80): 'I80'
    }
    df['condition'] = df.apply(lambda row:
conditions_map.get((row['mode[str]'], row['percentage']), ''),
axis=1)

    return df


def x_px_to_mm(px, total_px=xlimits_px[1],
total_mm=xlimits_mm[1]):
    return round((px * total_mm) / total_px, 2)


def y_px_to_mm(px, total_px=ylimits_px[1],
total_mm=ylimits_mm[1]):
    return round((px * total_mm) / total_px, 2)


def set_to_mm():
    df['goal point x[mm]'] = x_px_to_mm(df['goal point x[px]'])
    df['goal point y[mm]'] = y_px_to_mm(df['goal point y[px]'])
    df['guessed point x[mm]'] = x_px_to_mm(df['guessed point
x[px]'])
    df['guessed point y[mm]'] = y_px_to_mm(df['guessed point
y[px]'])
```

```
df = read_data('data')
set_to_mm()

x = df['goal point x[mm]'].to_numpy()
y = df['goal point y[mm]'].to_numpy()
gx = df['guessed point x[mm]'].to_numpy()
gy = df['guessed point y[mm]'].to_numpy()
conds = df['condition'].to_numpy()

indices_G = np.where(conds == "G")[0]
Gx = x[indices_G]
Gy = y[indices_G]
Ggx = gx[indices_G]
Ggy = gy[indices_G]

indices_P = np.where((conds == "P60") | (conds == "P80"))[0]
Px = x[indices_P]
Py = y[indices_P]
Pgx = gx[indices_P]
Pgy = gy[indices_P]

indices_I = np.where((conds == "I60") | (conds == "I80"))[0]
Ix = x[indices_I]
Iy = y[indices_I]
Igx = gx[indices_I]
Igy = gy[indices_I]

indices_GP = np.where((conds == "GP60") | (conds == "GP80"))[0]
GPx = x[indices_GP]
GPy = y[indices_GP]
GPgx = gx[indices_GP]
```

```
GPgy = gy[indices_GP]

AccG = np.average(np.linalg.norm(np.array([Gx, Gy]).T -
np.array([Ggx, Ggy]).T, axis=1))
AccP = np.average(np.linalg.norm(np.array([Px, Py]).T -
np.array([Pgx, Pgy]).T, axis=1))
AccGP = np.average(np.linalg.norm(np.array([GPx, GPy]).T -
np.array([GPgx, GPgy]).T, axis=1))
AccI = np.average(np.linalg.norm(np.array([Ix, Iy]).T -
np.array([Igx, Igy]).T, axis=1))

# >>> AccGP
# 83.68090905328367
# >>> AccG
# 103.09136430357988
# >>> AccP
# 102.33974571963797
# >>>



indices_P60 = np.where(conds == "P60")[0]
P60x = x[indices_P60]
P60y = y[indices_P60]
P60gx = gx[indices_P60]
P60gy = gy[indices_P60]
AccP60 = np.average(np.linalg.norm(np.array([P60x, P60y]).T -
np.array([P60gx, P60gy]).T, axis=1))

indices_P80 = np.where(conds == "P80")[0]
P80x = x[indices_P80]
P80y = y[indices_P80]
P80gx = gx[indices_P80]
P80gy = gy[indices_P80]
```

```
AccP80 = np.average(np.linalg.norm(np.array([P80x, P80y]).T -
np.array([P80gx, P80gy]).T, axis=1))

# >>> AccP60
# 120.54267709656533
# >>> AccP80
# 84.27193684497837

indices_GP60 = np.where(conds == "GP60")[0]
GP60x = x[indices_GP60]
GP60y = y[indices_GP60]
GP60gx = gx[indices_GP60]
GP60gy = gy[indices_GP60]
AccGP60 = np.average(np.linalg.norm(np.array([GP60x, GP60y]).T -
np.array([GP60gx, GP60gy]).T, axis=1))

indices_GP80 = np.where(conds == "P80")[0]
GP80x = x[indices_GP80]
GP80y = y[indices_GP80]
GP80gx = gx[indices_GP80]
GP80gy = gy[indices_GP80]
AccGP80 = np.average(np.linalg.norm(np.array([GP80x, GP80y]).T -
np.array([GP80gx, GP80gy]).T, axis=1))

# >>> AccGP60
# 93.59049189891167
# >>> AccGP80
# 84.27193684497837

rt = df['reaction time[s]'].to_numpy()
Grt = rt[indices_G]
Prt = rt[indices_P]
GPrt = rt[indices_GP]
```

```
Grt.mean()
Prt.mean()
GPrt.mean()


# >>> Grt.mean()
# 0.6415005888846214
# >>> Prt.mean()
# 0.9841778174304404
# >>> GPrt.mean()
# 0.9001172110398695

# indices_I = np.where((conds == "I60") | (conds == "I80"))[0]
# Irt = rt[indices_I]
#
# Irt.mean()
## 0.9238603420079033

def join(A, IA, B, IB):
    # Create a dictionary for quick lookup of B by IB
    ib_dict = {key: val for key, val in zip(IB, B)}

    # Find common keys
    common_keys = np.intersect1d(IA, IB)

    # Build C
    C = []
    for i, ia in enumerate(IA):
        if ia in ib_dict:
            C.append([ia, A[i], ib_dict[ia]])

    return C
```

```
# NARS
# attitude towards robot

ids = np.array([v.split('-')[0] for v in
df['participant_ID'].to_numpy()])
quest = pd.read_csv(os.path.join("questionnaries",
"negatt2robot.csv"), sep=';', header=None)
qids = quest[0].to_numpy()
scores = quest[[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]].to_numpy()
avgscores = np.average(scores, axis=1)
accs = np.linalg.norm(np.array([x, y]).T - np.array([gx, gy]).T,
axis=1)

# >>> avgscores.min()
# 1.8333333333333333
# >>> avgscores.max()
# 5.0
# >>> avgscores.mean()
# 3.478395061728395

d = join(accs, ids, avgscores, qids)
dids = np.array([v[0] for v in d])
daccs = np.array([v[1] for v in d])
davgscores = np.array([v[2] for v in d])
ddislikes = np.where(davgscores > avgscores.mean())[0]
dlikes = np.where(davgscores <= avgscores.mean())[0]
daccdislikes = np.average(daccs[ddislikes])
dacclikes = np.average(daccs[dlikes])

daccdislikes
dacclikes
```

```
# >>> daccdislikes
# 93.70260429156757
# >>> dacclikes
# 94.80662739313901

correlation = np.corrcoef(daccs, davgscores)[0, 1]
# >>> correlation
# 0.12896669612271686

# NASA
# task load
nasa = pd.read_csv(os.path.join("questionnaries", "taskload.csv"),
sep=';', header=None)
nids = nasa[0]
frustrations = nasa[6]

acc_dict = defaultdict(list)
for id_, acc in zip(ids, accs):
    acc_dict[id_].append(acc)

frustration_dict = {}
for id, frustration in zip(nids, frustrations):
    frustration_dict[id] = frustration

avg_dict = {k: np.mean(v) for k, v in acc_dict.items()}

acc_list = []
frustration_list = []
for id in avg_dict:
    if id in frustration_dict:
        acc_list.append(avg_dict[id])
        frustration_list.append(frustration_dict[id])
```

```
correlation = np.corrcoef(np.array(acc_list),
np.array(frustration_list))[0, 1]
correlation
# >>> correlation
# 0.20718332324567967


#
```