LLM-based Social Robot for Pleasant Supervised Visit in Security Waiting Rooms

Iterative Design of a Social Robot Supporting Visitor Mood and Staff Situation Awareness in a Security Waiting Room

Jean-Paul Smit

Delft University of Technology

TU Delft

LLM-based Social Robot for Pleasant Supervised Visit in Security Waiting Rooms

Iterative Design of a Social Robot Supporting Visitor Mood and Staff Situation Awareness in a Security Waiting Room

by

Jean-Paul Smit

Jean-Paul

Smit

Instructor:	Prof. Dr. M.A. (Mark) Neerincx
Daily Supervisor:	Ir. A.C. (Anne) Kemmeren
Thesis Committee:	Prof. Dr. M.A. (Mark) Neerincx, Ir. A.C. (Anne) Kemmeren
	Dr. N. (Nazli) Cila, Dr. C. (Clare) Shelley-Egan
Project Duration:	March, 2025 - Februari, 2026
Faculty:	Faculty of Computer Science, Delft

TU Delft TNO innovation for life

Abstract

Visitors of public security facilities (i.e. detention centres, courthouses, penitentiaries and customs) often experience anxiousness and low mood during their time in the waiting room, while staff have limited capacity to attend to visitors. To address this, we explore the design of a social robot that engages visitors through socially appropriate interactions intended to support a positive mood, while simultaneously providing staff with remote situation awareness.

Simultaneously, social robots have been deployed as reception agents and mood-intervention tools in public settings (e.g., healthcare[1] and retail[2]). Usually, social robots in public buildings handle welcoming and giving information; this thesis will test what happens when a reception robot also provides visitors a mood-intervention activity. A social robot could therefore provide a friendlier atmosphere in the emotionally fraught visitor area of public security constrained facilities.

Here we identify three complementary robot roles – receptionist, storyteller, and playmate – combined with a monitoring function for staff. Receptionist and storyteller roles were lab-evaluated using a local LLM powered social robot as a mood-intervention agent in two iterations with a Before-After Control-Impact (BACI) study. In the lab, the design was implemented with a Softbank Pepper-robot, a pre-post study indicated that interaction effectiveness depends on turn-taking level: high turn-taking activities produce larger immediate mood gains than low turn-taking storytelling. The playmate role was field-tested in a security-constrained public service waiting room. Beyond visitor surveys, semi-structured interviews explored staff acceptance. In the real-world evaluation, the design was implemented with an Enchanted Tools Miroki robot. The results indicate that visitors particularly value the playmate role, the guessing game agent was well accepted: 89% of visitors self-reported that the robot improved their visit ($t(9) = 4.0, p < .005$). Staff interviews underline acceptance rests on sociotechnical framing, perceived operational usefulness and minimizing of operational overhead.

The results advance understanding of social robots in high-consequence domains and other affectively complex (anxiety-laden) public places, suggesting that rapid, reciprocal interaction designs improve mood-intervention outcomes for security constrained facility visitors. Overall, the study findings inform the iterative design of social robots for sensitive, high-security public environments.

Our findings suggest further elaborating the waiting room robot with a (1) time-aware conversation system to offer activities suited to visitors' time-frames and (2) multilingual functionality for visitors without Dutch proficiency. Additionally, the other areas of the security constrained facility provide possibilities where robots may help reduce staff burden; (3) as a multilingual interpreter and as a (4) supporting buddy to reflect on accidents during visit to the facility.

Keywords: *Social Robots, Reception Robots, Mood Intervention, Security Waiting Room, Human Robot Interaction.*

Preface

This thesis is the result of several months of research, analysis, development, testing, (and more development ...), and the support of many people. I am grateful to everyone who contributed, be it intellectually, practically, or emotionally to this piece of academic prose.

First and foremost, special thanks to Prof. Dr. Mark Neerincx, my main supervisor at the University, for his encouragement and guidance throughout the process. I owe heartfelt gratitude to my parents Ineke and Willem-Jan Smit for their unwavering support and belief in me. I thank the thesis commmittee members Clare Shelley-Egan, PhD and Nazli Cila, PhD for their attention and consideration on this research. I am grateful to Anne Kemmeren, Birgit van der Stigchel, Wouter Meijer and Niels Cornelissen for their valuable feedback, insightful discussions, and constructive critique, which greatly improved the quality of this thesis. I thank TNO for providing resources and an inspiring research environment that made this project possible, TU Delft for offering a rigorous and stimulating programme in Computer Science and Leiden University (where I completed my BSc in Chinese Studies) for teaching me how to apply theoretical frameworks to real-world issues, connect multiple disciplines, and conduct fieldwork. My sincere thanks also goes to the team at the public security facility for putting care into action through their work, for the insights into their work they offered. They made a lasting impression and their commitment to care underscores that people-work must be bolstered — not replaced — to honour those who do it in the coming decades. Finally, I extend my sincere appreciation to all study participants whose time and involvement were essential to this research. To everyone who contributed in ways both large and small, thank you.

I acknowledge that while the ideation process (found in the Appendices: brainstorming, validating system prompts, generating system output) of this research has been proportionally LLM-assisted, a substantial part of the writing in the body of this thesis has come to be with limited use of large-language models. Personally, I believe that only inputting a simple prompt into a chatbot and then publishing its copy-pasted output with minimal redacting cannot be called true authorship, and would violate (my) academic integrity. So for most of the thesis writing, synonyms and definitions were found by use of the Merriam-Webster Thesaurus and Dictionary[1] or by simple web search [2]. For spell-checking (TU Delft uses British spelling) I relied upon Overleaf's native spell-checker and human proof-readers. If there is an em-dash (—) it is there because I appreciate them, not because I copy pasted them haphazardly from a chatbot.

Jean-Paul Smit
Delft, March 2026

[1] https://www.merriam-webster.com
[2] https://duckduckgo.com/

Contents

Nomenclature

Abbreviations

Abbreviation	Definition
AGC	Automatic Gain Control
AI	Artificial Intelligence
ASR	Automatic Speech Recognition
BACI	Before-After-Control-Impact (design)
CAG	Cache Augmented Generation
ET	Enchanted Tools
GUI	Graphical User Interface
HCI	Human-Computer Interaction
HRI	Human-Robot Interaction
LLM	Large Language Model
RMS	Root Mean Square
RAG	Retrieval-Augmented Generation
SCE	Socio-Cognitive Engineering
SCM	Stereotype-Content Model
STT	Speech-to-Text
TNO	Netherlands Organisation for Applied Scientific Research
TTS	Text-to-Speech
VAD	Voice Activity Detection
VSD	Value-Sensitive Design

1 Introduction

Security-constrained waiting rooms in public service facilities are often characterized by strict procedures, surveillance, and limited opportunities for staff to attend to visitors' emotional needs. Visitors in such environments may experience anxiety or low affect, particularly when the setting is unfamiliar and waiting periods are experienced as awkward or uncomfortable. These dynamics have been reported in waiting rooms such as airport security, border control, and court waiting areas, where prolonged waiting, uncertainty, and dependency are associated with negative affect (e.g. [3, 4, 5]). Penitentiary visitation provides a salient example of this broader class of waiting rooms, where visitors may experience relatively low affect; the environment and procedures are unfamiliar to many and can feel uncomfortable, sometimes resulting in visitors feeling stronger negative affect than the incarcerated people they are visiting [6]. Furthermore, visitation typically requires strict adherence to appointment schedules, removal of mobile devices, and waiting in a controlled access space for up to 30 minutes before continuation of their visit. This results in visitors spending a substantial amount of time in a liminal, security constrained waiting environment, which may further exacerbate negative mood and discomfort.

Security staff members face an important, yet challenging professional environment that may also contribute to work allocation challenges. For example, staff in the penitentiary play a crucial role in sustaining a rehabilitative environment for incarcerated people while also attending to visitors' needs. However, the tasks for these important workers are not popular, because the work does not come without dangerous situations [7]. Mainly that is because officers deal with individuals who have some record of harassment or violence: safety of everyone involved is essential: officers are expected to perform body checks on inmates, and they are actively confronted with crisis events of varying magnitudes such as: escapes, withdrawal, suicide or hostages [1]. For quite some time already these phenomena have contributed to the struggle of the Dutch security workforce with numerous vacancies, a high absenteeism rate, and a significant turnover rate among its staff [8]. Integrating social robots into security facilities, while prioritizing safety, requires changing officers' work environment and carefully addressing adversarial behaviour to ensure robots effectively support existing processes.

To support public security facilities, studies should investigate how assistive robots can enhance specific care tasks performed by staff (e.g., visitor info, counselling, inspections, control). Human–Robot Interaction offers guidance on designing such social robots (e.g. [9]), but few studies target the high-risk domain like penitentiaries. Siegmann and Bendel [10] suggest visitor-area robots could provide information and act as contact points for specific issues (as a sort of reception agent); social robots could ease staffing challenges (e.g., providing alternatives when female staff are required for women's body checks or making security monitoring more pleasant), and reducing visitors' anxiety through entertainment (mood intervention). While social robots appear in healthcare, retail, and libraries, their use in security-constrained waiting rooms remains largely emergent.

Visiting a security facility with waiting room generally follows a fixed sequence: visitors are welcomed,

[1] https://www.werkenbijdji.nl/vacature/beveiliging/complexbeveiliger-43817. accessed on 2-4-2025

pass a detection gate or bag search, and are escorted to a waiting room until the visitation hour begins, when they are collected for the visit. Visitors, especially on first visits, often report nervousness and anxiety. Inmates mood levels are usually not high [11], but show peaks at visitation hours [12]. A recent Dutch prison found that visitors' positive emotions decline after visits and can fall below inmates' positive emotions [6]. These observations suggest a clear need to explore how a robot can intervene in the mood for visitors and reduce stress during and after waiting for the visitation hour.

Building on suggestions from Siegmann and Bendel [10], we propose researching a social robot deployed specifically in the visitor area to address those needs. The robot would welcome and escort visitors to the waiting room, act as a clearly identifiable information and contact point, and provide entertainment and mood-contagion interventions to reduce anxiety [13]. We explore the ideas (for social robots in the security constrained visitor area) brought up by Siegmann and Bendel [10]: (1) information delivery; (2) mood intervention, and (3) support for staff by monitoring.

1.1. Research Objectives

This thesis will present the design of a social robot supporting the task of reception and mood improvement for visitors. It aims to initiate an innovation to address the challenges found by the field investigation at the security-constrained waiting room while exploring the further application of a reception robot in the security waiting room.

The research pursues two main research objectives. Firstly, it will investigate how a robot could improve mood in the security facility by creating a pleasant atmosphere through playing an informative, anxiety reduction role in receiving and guiding groups of visitors. Secondly, it will also investigate whether the support of the robot can decrease the task load of security staff and if the work becomes more meaningful with it, while keeping the current safety and privacy.

The research questions of this paper are as follows:

1. How can a robot improve the mood of a visitor of a security facility, securely for its operational domain, by using appropriate activities from a receptionist, storyteller or playmate perspective?
2. What operational functionalities can the robot equip to take good care of visitors upon arrival (e.g., welcoming, informing, keeping an eye on things) and while they are waiting within the security facility (e.g., mood intervention activities mentioned in RQ1)?

1.2. Thesis Outline

The following chapter describes the state-of-the-art of reception robots and the current challenges of technology in security facilities. That chapter forms the basis for the specification of the Design Rationale in chapter 3, including an initial stakeholder analysis deducted from meetings and visits to security-constrained facilities. The general prototype (section 3.7) was then developed iteratively, following from an evaluation in a university lab in chapter 4, and a formative field evaluation in chapter 5. Findings and interpretations will be discussed in chapter 6. Supporting detailed information such as the robot's LLM prompts, briefing and consent forms can be found in the Appendix A.1.

2
Background

This section has three parts that go from existing situations towards a projected enhancement of security-constrained waiting rooms. First, it will discuss social robots in waiting rooms. Then, it will discuss the characteristics of the public waiting room area of security-constrained facilities. From there it elaborates on the complexities of feelings that visitors may be experiencing. Lastly the state-of-the-art of robots, that may enhance the mood of visitors, while helping security staff are being summarized.

2.1. Social Robots in Waiting Rooms

Social robots have been studied across public waiting spaces with relatively low risks (e.g. hospitals [1], dentists [14]) and medium risk (e.g. psychiatric clinics[15], public airports [5]) showing they can lower stress or improve the experience of a public waiting area. Public waiting areas are particular as they pose an initial stage of contact at public buildings, and they are situated in liminal areas where visitors pass through and have an informational need, which is why usually reception robots can be found in these spaces. They will have the functionalities of welcoming visitors, providing information, guiding the way, identification and appointment management for visitors. Some will manage appointments until they can remind or call in a visitor for their appointment to start. Reception robots can be an efficient and useful tool to support staff in a public waiting room by relieving some traditional receptionist tasks, even facilitating some friendly chit-chat. However, current research does not touch upon mood intervention for visitors in a security constrained waiting room.

2.1.1. Social Robotics

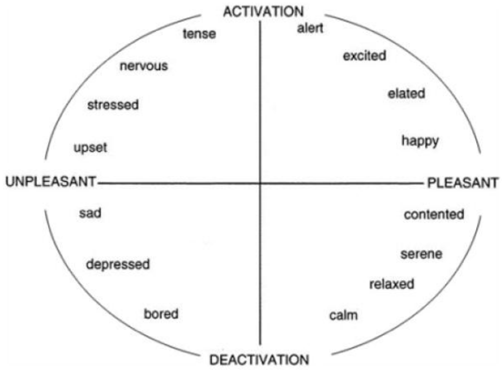
Social robots are robots which socially interact with humans through verbal and non-verbal manners to communicate and achieve objectives specific to their applications within society. When a conversation system has conversational memory it can be called a conversational agent, when embodied (virtually or physically) an Embodied Conversational Agent, or when physically embodied; a robot [16, 17]. The advent of Large Language Models (LLMs) has rapidly increased development and applications of social robotics, providing strong models for rich dialogues and communication for robots [18]. Large language models appear to have the abilities of generation and even reasoning. The capacity of LLMs has mainly been focused on text-based tasks such as writing, coding and translation. LLMs can also be embodied by agents with multiple modalities or even anthropomorphic features.

Embodiment

Social robots must both perform tasks and interact naturally with people; a robot's appearance, its embodiment, influences users' expectations, comfort, and the nature of interactions. Researchers apply the Stereotype Content Model, which characterizes social judgments along two dimensions—Warmth (friendly, approachable) and Competence (capable, effective)—to predict user preferences for different robot forms [19]. Figure 2.1a shows the model's four quadrants and their typical social judgments. This framework explains why certain embodiments are preferred for particular roles: people generally favour warm-looking robots for socially assistive roles and robots that signal competence for instrumental, task-focused roles [20]. Notably, robots perceived as low in competence can still be desirable if they are high in warmth (for example, ‘cute’ robots). Conversely, robots that appear highly competent but low in warmth, such as the Spot robot, may be perceived as eerie yet remain valuable for tasks like life-saving humanitarian missions or defence operations [21].

	Competence	
Warmth	Low	High
High	**Quadrant 2** **Paternalistic prejudice** low status, not competitive pity, sympathy e.g., elderly and disabled people, housewives	**Quadrant 1** **Admiration** high status, not competitive pride, admiration e.g., in-group, close allies
Low	**Quadrant 3** **Contemptuous prejudice** low status, competitive contempt, disgust, anger, resentment e.g., welfare recipients, poor people	**Quadrant 4** **Envious prejudice** high status, competitive envy, jealousy e.g., rich people, feminists

(a) Stereotype Content Model (warmth × competence) characterising social judgments of others [19]



(b) Affect Circumflex (valence × arousal) mapping momentary emotional states [22]

Figure 2.1: Two figures illustrating models for interpersonal perception (e.g. human reactions to robot embodiment) and for quantifying users’ internal mood.

- Humanoid robots, a popular medium for conversational robots, resemble people but typically avoid exact human likeness to prevent the uncanny valley: an unsettling effect some observers feel toward almost human forms. Humanoids often support rich verbal channels: expressive speech, synchronized lip movement, and gesture, which make them suitable for reception, information, and conversational roles (for example, SoftBank Robotics’ Pepper see Figure 2.2c).
- Zoomorphic robots, popular for security applications or the mood intervention domain, take animal forms. Examples include the task-oriented Boston Dynamics’ Spot (see Figure 2.2a) and service-oriented Sony’s Aibo (Figure 2.2b). These designs often communicate primarily through non-verbal channels, body posture, movement, and simple sounds, and can invite touch, play, and therapeutic interaction. Smaller, petlike robots like the Aibo may be used in healthcare and elder care as comforting companions, whereas the Boston Dynamics’ Spot has been deployed in security and institutional settings where its presence can feel intimidating, or even chilling [21].
- There is also an emerging variant of hybrid designs that blend human and animal cues stylized characters that are neither strictly humanoid nor purely zoomorphic, like Enchanted Tools’ ‘Miroki’ (Figure: 2.2d). These fantasy-inspired forms combine the communicative clarity of anthropomorphic faces with the approachable, playful affordances of pet-like bodies, yielding benefits from both archetypes.

We will explore Pepper and Miroki as candidate platforms for implementing our conversational system.

Reception Robots

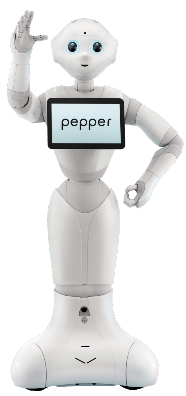
The field of reception robots is experienced with the tasks described by [10] for visitor areas (welcoming and informing) where social robotics are use for greeting, answering questions, providing directions [26, 27, 28, 29, 30]. Receptionists are important because they shape visitors’ perceptions of the facility[31], but reception robots may also handle guest check-in, data registration, and way-finding across sectors such as healthcare, retail, home, and education [26]. Before the advent of LLMs, that happened with Finite state machines or rule based logic [32], but since LLMs run on robots we use information bases and RAG. Many information-providing robots use Retrieval-Augmented Generation (introduced by [33]) to ground LLM outputs in up-to-date sources and reduce hallucinations. RAG-based reception systems therefore combine a retriever (external documents/web results) with a generator, so responses are factual and verifiable [34].

LLMs in the Security Domain

High risk domains see itself as lacking behind with technology but is often a test-bed for it actually and one of the things it might try out is robots[35]. Robots are nice for high risk domains where work is

(a) Boston Dynamics Spot [23]

(b) Sony Aibo: 2018 version [24]



(c) Softbank Robotics Pepper [25]

(d) Enchanted Tools Mirokai

Figure 2.2: The zoomorphic quadrupedal Spot, zoomorphic Aibo, the anthropomorphic Pepper and the cartoon-figure-inspired Miroki

dangerous or dirty for humans. The high-consequence domain in which security facilities are operating, as we saw in sections before, is like one of those domains, but it also requires attentive handling of user data and professional data.

Current issues with data privacy and security practices highlight the importance of transparent platforms. Since the breakthrough of small, quantized models like DeepSeek, local LLMs can run on relatively small devices. Quantized models exhibit higher tendency to produce false or invented facts; mitigating measures include conservative decoding (set temperature to 0), strict prompting that constrains output scope. Firstly because the local model does not require outside connection to an external provider of services, there is less decoupling of tasks and less risk of unwanted interference from outside. Secondly, the local LLM only has access to a set of answers and data that is set, inferred or downloaded on its local machine. These applications mitigate risks from outside and may provide data privacy for its users, which is especially important for public service facilities where security is constrained.

To mitigate hallucinations the LLM should run locally on the robot, disconnected from any internet access. Using a cloud/external API based LLM in a security constrained facility poses unacceptable security and privacy risks. Local LLM runtimes avoid external network exposure but historically required large, costly hardware; recent advances in model optimization and quantization make on-premise deployment feasible on much smaller machines with reduced perceived performance: answers may be factually incorrect way more. That perceived performance gap is commonly closed with Retrieval Augmented Generation (RAG) that injects up-to-date web search results for the context, but security wise, internet access is inappropriate in most high risk environments. Therefore, we use the local version of RAG called CAG; Cache augmented generation [36]. Accordingly, RAG must be restricted to local knowledge data: an on-device, curated corpus.

Adversarial and malicious prompting to LLMs requires a careful approach within the high risk domain. While many models already employ moderation into their models, one can improve the robustness by sandboxing and context restriction of the models. That means deploying the LLM in a controlled execution environment (sandbox) with fixed system prompts, redaction of sensitive user inputs, no outbound API/shell/actuator access, and tightly limited tool-calling privileges. Combine that with safety filters or deterministic rule blocks for forbidden topics (e.g., weapon fabrication questions by visitors or making the robot target staff members).

2.2. Security-Constrained Visitation

Visiting security-constrained areas, such as penitentiaries, is emotionally charged: visitors often arrive in a markedly lower mood [6]. A visit to a security waiting room starts with arrival: the visitor arrives at the entrance where they are welcomed and asked to hand in any prohibited items (e.g. smartphones), and then passes through 'security': a detection gate to confirm no prohibited items are smuggled past the entrance. After that, visitors take a seat in the waiting room with others until a staff member escorts them to the visitor room. The waiting period can be particularly distressing, because visitors are confined to a monitored space, face uncertainty about timing, do not have mobile devices at their disposal during the waiting time, and in some contexts (e.g. penitentiaries, police stations, immigration services) are continuously reminded of the institution's punitive context, all of which amplify anticipatory stress and emotional strain before the visit starts. Formal visiting procedures (i.e. security checks, controlled movement, and time-limited interactions) can intensify that strain.

Security-constrained visits offer valuable opportunities but entail a distinct set of regulatory and practical challenges. Among cultures, one of the most positive affective moments in the experience of incarcerated people is the penitentiary visit [37]. The visitation is part of a humane incarceration process and has in that sense a goal, but other than that it has been shown to support the rehabilitation process and reduce recidivism [38]. Generally, visits must be scheduled in advance, and depending on the facility, visitors may be permitted a brief hug and kiss with their loved ones. In contact visits, individuals can sit together, usually at a small table, for a limited period. It is interesting to note that whereas across countries most incarcerated people are always male, most visitors are women [6] [39]. Strict regulations are in place to prevent contraband, and visitation can be categorized as either non-contact or contact. Apart from global similarities, some things beg to differ across countries.

2.2.1. Visitation Challenges

For security staff, visits are a double-edged sword: while they play a vital role in upholding the humane rehabilitation of incarcerated people, they also create environments of increased adverse behaviour [12], where contraband can be introduced [40], thereby complicating efforts to maintain security and order within the facility. Contraband can be described as having drugs, illegal substances or items in possession (such as smartphones). The issue is pressing, for example in the penitentiary sector, an analysis on 3885 Dutch incarcerated people shows that contraband smuggling is increased by 63% for incarcerated people with visits compared to incarcerated people without visitors [40]. Seen the risks that come with visitation, some argue that remote visitation, similar to the online visitation of quarantined people [41], can eliminate the challenges of in-person visitation. But visitors are frustrated with the disadvantages and lacks of remote visitation [42] because video communication formalizes conversations [43], limits visual cues, and hinders trust-building, disrupts natural communication, increases social distance, and leads to shorter, less interactive and intimate exchanges [44]. Both supplemental and as-replacement video visitation are thus linked to adverse effects on rehabilitation and therefore higher recidivism [45]. It is clear that remote visitation is inferior to the in-person visitation for reasons that mainly have to do with affect.

2.2.2. Visitation and Mood

In the context of emotional landscape of security facilities such as the penitentiary, the visitation is universally self-reported by incarcerated people as a positive point for incarcerated people [37]. The security waiting room is an interesting landscape of different emotions which have been studied extensively within the incarcerated populations of penitentiaries [11]. Sykes [46]'s theory of the 'Pains of Imprisonment' describe the psychological and emotional suffering experienced by incarcerated people due to loss of freedom, social isolation, disrupted relationships, and the harsh realities of life in the facility. Others like Fredriksson [47] further attribute these 'pains' to violence in the forms of the loss of the self, loss of subjectivity, haunting of the past with the present (this haunting is especially pronounced in certain visits, more on that in the next paragraph). Inmates have the infamous 'inmate code' and incarcerated people generally have higher levels of anger compared to the general population outside the security facility. In this bleak environment, the visitation is a beacon for incarcerated people. According to a 2022 study in Israeli security facilities, incarcerated people levels of hostility and anger decreased following visits [12].

Yet the visitation has a heterogenous effect on incarcerated people and so it may also foster negative affect and hostility, depending on the visitor and topics[6], whereas visitors mostly have negative affective experience after visits. The negative experiences surrounding the visit can be explained through a hauntological lens, as the outside world is brought in with the visitors, which reminds the incarcerated person of the past [48]. But coming to terms with the outside world is necessary for rehabilitation, which is why the situation is more complex than the framework of Derrida [49]: de Jong et al. [6] pioneered an academic effort to map incarcerated persons and visitor feelings to topics discussed during visitation hours. Their findings show that emotional topics, such as asking how one feels during the visitation hour are related to positive experiences, and problem-solving topics, such as finances, are related to negative experiences among incarcerated people. Although the visit is self-reportedly to be a rather positive experience for incarcerated people, the effects on visitors also need consideration, because de Jong et al. [6] finds that visitors have more negative affective experiences than incarcerated people.

Another theory which explains the negative feelings of visitors of security facilities such as the penitentiary is the concept of 'secondary prisonisation'. As Comfort [39] explains, visitors can absorb some of the environment's stress and culture, making their visit, and life after the visit, less comfortable. Her auto-ethnographic study explores how almost solely female visitors repeatedly undergo this process which unfolds in three parts: pre-visit visitors prepare for visit, during visit they submit to the rules and gaze of the security facility, and they adjust their behaviour to retain admittance to the visitation. The visitors explain that they continuously redefine their integrity, both socially and physically, to prevent a certain degradation of identity that is expected in the facility under the guise of 'security' practices [50]. (legally free people from outside have to be transformed into non-threatening entities that are constantly under the surveillance, ánd gaze of the facility with its rules and punishment)) This means the 'Pains of imprisonment' [46] of the incarcerated person are extended to the realm of their visitors. The research of de Jong et al. [6] dates from 2022, and as of now there has not been any reported integration of social robotics to change the negative affective experiences of visitors.

Lastly, the objects within the ‘liminal space’ of the waiting room before the visitation impacts the mood of the visitors. Waiting areas in public sectors can be a place that increases stress or anxiety, but also create a peaceful feeling of elevation in visitors. By using the right media, such as certain types of imagery that promote connection with others and nature, waiting rooms can improve their visitors’ mood to achieve a feeling of ‘elation’[51]. In contrast to the health care waiting room, the security constrained waiting room is more emotionally challenging, because personal objects such as smartphones are left at the entrance, which might bore visitors even more. This might not seem a problem, but boredom is not harmless, people who are prone to it may find it relatively stressful, and chronic boredom is associated with a decrease in overall lifespan. The choice of elements in the waiting room have an impact on the mental health of visitors.

2.3. Mood Intervention Activities

Designing robots to intervene effectively in people’s moods requires coupling a socially engaging character with reliable interactional competence. Successful mood support robots therefore must (1) present a friendly, relatable persona and (2) demonstrate robust abilities to perceive, interpret, and respond to human affective states. Affective states are commonly described along two orthogonal dimensions; valence (pleasant to unpleasant) and arousal (activated to deactivated) [22].

2.3.1. Mood versus Emotion

While systemic factors create an emotionally taxing landscape in security constrained waiting rooms, environmental adjustments can effectively shift visitor mood. In Human-Robot Interaction, emotions are described as being intense, short-lived reactions to specific objects (e.g., the grief of a relative’s incarceration), whereas mood is a diffuse, low-intensity background state [52]. Especially in the security-constrained domain, measuring emotion is impractical and risks offering mere “technological fixes” to systemic social problems [53]. Structural issues like power asymmetry, uncertainty, and secondary prisonisation[39] impose heavy emotional burdens that design cannot resolve. However, some avoidable low mood can be mitigated in waiting rooms, for example removing distressing media from waiting rooms can improve mood[51]. Consequently, interventions in mood is sensitive to environmental ‘small wins’ and can display improvements in the immediate human experience without over-claiming to overhaul the underlying structural complexity.

2.3.2. Mood Intervention Activities

Mood-intervention via distraction or brief social activities has empirical support in healthcare: short, engaging interactions reduce anxiety in clinical contexts (e.g., pre-vaccination[1], blood draw [54], in dental procedures[14], dementia care [55]) and that suggests mood benefits when adapted to other high-stress waiting environments. In these contexts activity-based interventions (e.g., games, storytelling) serve as intervention regulatory strategies that redirect attention and provide emotional support, improving affective state.

Prior work suggests that a person-focused, enthusiastic robot that explicitly communicates its social capabilities elicits stronger social contact and greater user engagement [27, 56]. For mood-intervention effects to emerge, the robot must also demonstrate conversational fluency and task competence: poor turn-taking or conversational failures undermine engagement and can reverse intended mood benefits[57]. Consequently, fluency and responsiveness are central design requirements for any mood-intervention agent in applied settings.

Robotic emotion synthesis and perception draw on multiple modalities. Dialogue systems, verbal sentiment analysis and synthesis, and embodied signalling (facial expressions, gaze, gesture) are combined to portray affective states, often informed by frameworks such as the Facial Action Coding System [58] and Laban Movement Analysis for gesture and posture [59]. Humans naturally synchronize affective states with social partners — emotional contagion is a well-documented phenomenon in HRI [13] — and people can reliably recognise emotions expressed by robots [60]. Storytelling and narrative-rich interactions are particularly effective at promoting emotional resonance and contagion, making them a promising modality for mood-intervention robots [61].

2.3.3. Storytelling

Drawing on Brockington et al. [62], storytelling, whether delivered by humans or socially interactive robots, can increase oxytocin, reducing cortisol and reported pain, and promoting more positive affect; it also enhances empathy through character identification and fosters imaginative problem-solving, making narratives a useful tool for engaging multilingual children [63] and for social robots that share emotionally resonant human stories to strengthen feelings of connectedness like other media in waiting room contexts [51].

2.4. Rationale for Robot Choice

Robots are part of a larger ecosystem of smart technologies that are increasingly explored in high-security settings within managerial (e.g. surveillance, contraband detection, incident analytics, tracking, health care appointments management) and rehabilitative (e.g. educative, preparing for re-integration, e-banking, communication) use cases [64]. This research highlights visitor stress, time management, and staff concerns about safety and supervision, but contains no evaluations of social robots deployed in security constrained waiting rooms. Previous exploratory research [10] suggest visitor-area robots can provide information and act as contact points for specific issues (as a sort of reception agent). Social robots could ease staffing challenges (e.g. providing alternatives when female staff are required for women's body checks or making security monitoring more pleasant), and reduce visitors' anxiety through entertainment (mood intervention). It is clear that there are many potential roles for a social robot that can suit the security-constrained waiting room, though the effect of each role on visitor mood and staff operation are not yet clear.

To sum up, despite the growing interest in reception robots and technology, until now little exploration has been done on the role and impact of a reception robot in security-constrained waiting rooms. Adding a social robot in the security-constrained waiting room may affect the mood of visitors, and it is essential to study this effect.

Addressing this gap is important, because deploying a reception robot in the security constrained waiting room may have potential affective advantages to visitors and indirectly on the incarcerated people. It may especially decrease the workload of current security facility staff, improve the experience of visits of relatives, which may be mirrored by incarcerated people indirectly.

Contributions of this research are:

- Enhancing understanding of visitation experiences in security constrained waiting rooms.
- Expanding research on social robotics and human-robot interaction by developing and evaluating HRI use cases in this novel application domain.
- Advancing applied knowledge on deploying robots in high-consequence workplaces, with a case study on LLM-based robot interactions for emotion regulation during visits.

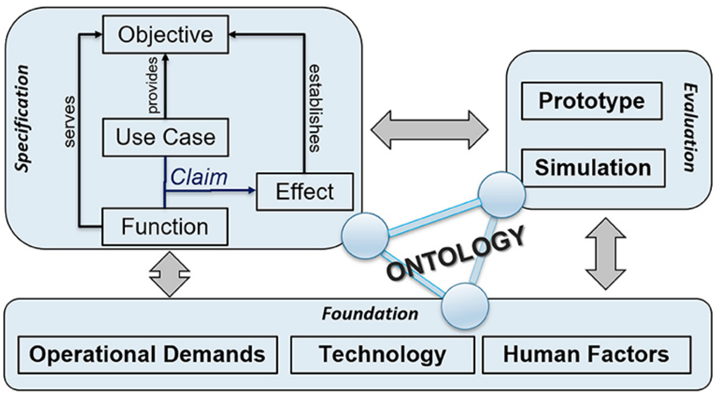


Figure 2.3: Overview of the Socio-Cognitive Engineering methodology from Neerincx et al. [65].

3 Design Rationale

Before deploying a robot, the context and scope of the project need to be channelled into a prototype, which is where this chapter comes in. To develop a proper framework for the robot's objectives and constraints, we compose the use cases, Human Robot Interaction (HRI) setting and system, which will be used to reach a viable specification for a robot in the security constrained waiting room.

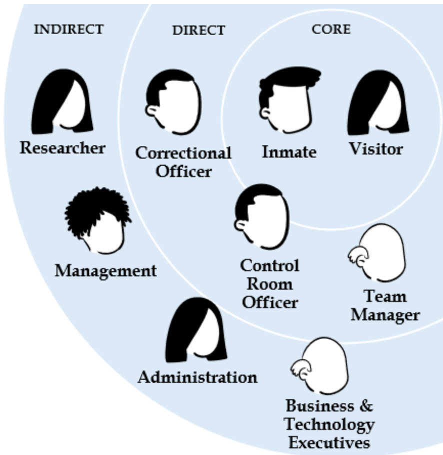


Figure 3.1: Stakeholder Circle Map. The Circles indicate how involved stakeholders are in the design process. There are three circles: direct-core (highest involvement), direct-personnel (medium involvement), and indirect (low involvement).

3.1. Approach

The design of the robot system will follow an iterative and incremental pattern, improving the requirements by use of prototyping or simulation, as described in the SCE (Socio- Cognitive Engineering) method: an approach used to design and develop complex systems in a step-by-step, evolutionary way [66]. It focuses on understanding how humans interact with technology by analysing operational needs, human factors, and technical aspects, see figure 2.3. This method starts with gathering and specifying system requirements through detailed use cases, requirements, and claim analyses, breaking down the overall system into smaller, manageable parts. Engineers then test these ideas using simulations or prototypes, allowing them to continuously refine and improve the design based on feedback. By iteratively making adjustments, the SCE method ensures that the final system is well-suited to its intended tasks and meets the needs of its users: the stakeholders at the core of the situation as visualised in figure 3.1.

3.2. Rationale

To reinforce the main objective of the operational demands: a robot may keep the work of security staff bearable and help visitors in the process, but it will not aim to replace staff. It is important to discover if a robot significantly improves the lives of staff and visitors, especially if it can avoid more burn out related events happening which will intensify workload due to more absences. In a holistic view on work ethics, the robot might make work more interesting, implying the need for meaning and fulfilment in one's job, which is as important as being paid a good salary or benefits [67].

To gather an understanding of the situation, we visited two security partners in March 2025. The current situation brings to light that the security staff carries out a vast amount of valuable tasks, like escorting visitors, strip searching individuals, monitoring situations but also giving care. In case there is labour allocation scarcity (suppose colleagues take an unexpected sick leave while there are no additional colleagues) the work tasks may get overwhelming. Similarly, visitors may deal with feelings of low mood or anxiety when confronted with a first-time visit, and they may have an information need according to the management during one of our meetings at the initial visits.

We sketch personas (see Table 3.1) based on the current situation in the security constrained waiting room workplace and visitation area to get a sense of its operational domains. The remainder of what follows from that information, the personas are sketched in Table 3.1 and form the basis of the scenarios in the following section.

Name (Age)	Description
Pierre (52)	Pierre is a staff member at the youth security constrained waiting room. Pierre enjoys his job and finds meaning in the care-giving to youth offenders, while keeping him and others safe. However, recently he has to take on a crushing amount of work because there are not enough colleagues to fill all staff positions. Because of that he does not have enough time to meet all the needs of visitors. As a result, Pierre may find his work load increasingly unfulfilling and risks overworking. He sometimes misses contraband brought in by visitors because his schedule is too busy.
Evelyn (39)	Evelyn comes to visit her son Sebastian (15) who is in the security constrained waiting room since last week after a property-related offence. She finds it devastating that her child is put in the facility. She is anxious about the visit, she has never set foot into a security facility like this before. She has quite some questions and feels frustrated that there is not an extensive amount of time to accompany her questions. To her, the lack of attention may seem as if there is not enough open or honest communication.
Jasmin (10)	Jasmin joins her mother Evelyn to the visitation hour. She is affected by her mother's uneasiness. Jasmin is also bored, she has to wait without her comforting iPad inside the waiting room, as a result, Jasmin starts to become uneasy which further upsets her mother Evelyn.
Sebastian (15)	Sebastian is incapacitated since last week due to a property-related offence. He looks forward to seeing his mother and he misses his sister Jasmin. He is still coming to understand the severity of the situation, but an emotionally positive visit may push him into the right direction.

Table 3.1: Four personas for the design of the security waiting room support

3.3. Scenarios

Scenarios are testable definitions of desires of envisioned situations. To define those, we first depict what the current situation looks like, as interpreted from meetings to security facilities and visits. Using that existing scenario we can imagine what a situation with the robot would look like and what it would be expected to do in these situations. Remember how in the personas (Table 3.1) we described four imaginary persons visiting, working or living inside the security constrained waiting room. Their descriptions show challenges in the current work.

The 'Problem scenario' in Table 3.2 describes how the staff currently operates around the visitation hour. Apart from that, we describe two more scenarios: the alternative and the design scenario. Alternative scenarios help to better come up with a design that actively participates in solving the problems. The alternative scenario in Table 3.2 uncovers three problems that may currently challenge the security staff

Problem scenario	Pierre ensures a safe, lawful environment for himself, visitors, and other individuals in the facility. He monitors people and situations to prevent smuggling of contraband, he is answering questions of visitors and helps visitors with their first encounters inside the security constrained waiting room. He is also accompanying people to the visiting area. During visitation hour he is paying attention to the group dynamics, to prevent escalation, smuggling of contraband and aggressive or hostile behaviour. Sometimes he cuts off the visitation hour prematurely due to prevention of one of these negative effects.
Alternative scenario	If there are not enough colleagues to fill all positions Pierre does not have the so-called 'bandwidth' to (P1) await Evelyn and Jasmin at the door to escort them to the waiting room and (P2) to answer all the questions of others. He may also not be occupied specifically with (P3) breaking awkward silences as to not 'rock the boat', which means that any anxiety or mood experienced by visitors is not intervened but also not escalated and (P4) he cannot oversee everything and may miss contraband. He is mainly focused on keeping everyone safe, giving priority to escorting people and de-escalation of any unwanted behaviour that he and his colleague evaluate as unsafe. He is not focused on answering questions from visitors.
Design scenario	Pierre employs the robot officer to assist him in answering questions of visitors, to break awkward silences and to gain better situational insights. Evelyn likes that the awkward silence is broken, she uses the robot to get answers to the questions that make her feel uneasy, and it brings her into a mood of relief and less frustration. For Jasmin, the robot with its zoomorphic qualities give her entertainment value that she misses after putting away her iPad. Sebastian is happy that his mother and sister are in a good mood, and it potentially affects his contemplative experiences afterwards.

Table 3.2: Three scenarios for the design of the security constrained waiting room support

and visitors:

P1. Reception — showing visitors their way around in the security constrained waiting room, *especially if it is their first visit*, in which case they may be anxious and/or have 'burning' questions.

P2. Information provision — answer the questions visitors may have about procedures or about the remaining waiting time before their appointment.

P3. Mood intervention — to *'break the ice'*, change the frame of mind of visitors into something less anxious, eventually improving their visit.

P4. Monitoring — to support the staff by observing visitor interactions, detecting irregularities or unusual behaviour, and flagging safety or welfare concerns (e.g., potential contraband attempts, aggressive interactions) so staff can intervene promptly and proportionately.

To support the staff, the robot may need to facilitate monitoring of the conversations and interactions. This is vital for a two-fold of reasons. Firstly to ensure that the system of the robot works properly (keeping a human in the loop). Secondly it parallels the current practice of security staff with monitoring devices existing inside security facilities, such as surveillance cameras. Thirdly because it will help them have an 'extra pair of eyes' to discover in case a situation escalates. Especially after an incident happened in the visitation area, the conversation history may be useful to perform a review and learn lessons from it to prevent future accidents.

We will come up with a design approaching the current challenges, where a robot is envisioned to be in use at the security constrained waiting room, working alongside the staff to fulfil the situational desires of visitors and staff.

3.4. Use Cases

Further refinement of the prototype foundation is done by breaking down the design scenario into multiple use cases. Use cases are specific action flows, which point out the requirements of an envisioned prototype design. Each use case also takes into account an alternative sequence, to ensure that the robot indeed may provide a solution to the problems of the scenario.

The problems in the design scenarios show that a security support robot should perform a main function of (1) reception agent, that it should (2) provide information to visitors and that it should (3) have a mood intervention effect.

Apart from the knowledge gathered from these scenarios, the robot should possess the appropriate functionalities for its domain for supporting the working conditions of the staff, i.e. *it should make their work easier not harder*. As described in the scenarios (Table 3.2), to support staff, the robot's conversations, especially in the context of the security constrained waiting room, should reduce unexpected and inappropriate content, possibly by means of moderation. The robot should also know of the sensitivity in its domain; it should know its constraints. To hold the robot accountable to erroneous responses, the conversation history should be temporarily saved and must be accessible to authorized staff members.

Figure 3.2: Use case 1 displayed in three steps.

Figure 3.3: Use case 2 displayed in three steps.

Furthermore, taking human factors knowledge into account, the literature review points out three functional wisdoms to factor into the design:

- baseline 1: a person-focused, enthusiastic robot which communicates its social capabilities makes better social contact and more engagement [27, 56, 68].
- baseline 2: the robot should be fluent/competent and engaging enough to have such a mood-intervention effect [69].

- hypothesized: mood-intervention or distraction activity can improve the mood of individuals in anxious or less cheerful settings like hospitals (i.e. [1, 70, 54]) and that effect may also extend into the security constrained waiting area of visitors.

Use Case 1	
Description	The use case describes the robot welcoming a visitor and providing information. The robot is recording the conversation history in a transcript accessible to authorized staff members.
Objective	(1) To satisfy the information need of visitors during visitation by answering their questions. (2) To support security staff by reducing cognitive load and providing happier visitors before the visitation hour starts.
Actors	Staff, Visitor, Robot
Pre-conditions	The robot is passively recording conversations. The robot is operational and stationed in the visitation waiting area. The visitor has passed security and is waiting for visitation. The robot has access to a library of information about the facility, the institution, the visitation regulations or the remaining waiting time. The robot is passively recording conversations.
Post-conditions	The visitor has had a fluent interaction with the robot. The visitor reports an improved visit The visitor reports improved mood or reduced anxiety. The interaction is logged for later review.
Action sequence	1. Visitor enters the waiting area. 2. Robot greets the visitor and initiates small talk. 3. Visitor responds and engages in dialogue. 4. Visitor asks questions 5. Robot answers questions 6. Visitor now has answers 7. Robot announces waiting time is over, and the visit is about to start. 8. Robot ends the conversation politely and logs the interaction. 9. Robot performs real-time transcription and records it in conversation history accessible to authorized staff.
Alternative sequence	1. The visitor does not engage with the robot. 2. The visitor's questions are not answered.
Requirements	**Requirement 1:** The robot shall welcome a visitor into the visitation area enthusiastically. **Requirement 2:** The robot must be able to initiate and maintain (for the entire waiting time duration) a natural conversation for with a visitor using everyday language that most visitors can understand. **Requirement 3:** The robot must be able to provide information to visitors. **Requirement 4:** The robot must be able to transcribe and record conversations.

Table 3.3: Use case one: one visitor comes to the waiting room with an information need.

Use Case 2	
Description	The use case describes the robot performing a mood-intervention activity where it shares a background story to engage the visitor emotionally. The robot is recording the conversation history in a transcript accessible to authorized staff members.
Objective	(1) To improve the mood and emotional well-being of visitors during visitation by a mood-intervention activity like storytelling or gaming from a robot with an expressive, enthusiastic character. (2) To support security staff by reducing cognitive load and providing happier visitors before the visitation hour starts.
Actors	Staff, Visitor, Robot
Pre-conditions	The robot is passively recording conversations. The robot is operational and stationed in the visitation waiting area. The visitor has passed security and is waiting for visitation. The robot has access to a library of mood-intervention activities, stories and conversation prompts.
Post-conditions	The visitor has had a fluent interaction with the robot. The visitor reports an improved visit The visitor reports improved mood or reduced anxiety. The interaction is logged for later review.
Action sequence	1. Visitor and robot engage in conversation. 2. Robot shares a background story designed to change mood. 3. Visitor responds and engages in dialogue. 4. Visitor thinks less about their anxiety and feels better 5. Robot announces waiting time is over and the visit is about to start. 6. The robot announces that the waiting time is over and the visit is about to begin 7. The visitor proceeds to the visitation area 8. The interaction is logged for later review.
Alternative sequence	1. Technical failure: the interaction shows unexpected responses, errors, repetitions and is not fluent 2. Technical failure: transcript or alert is not delivered.
Requirements	**Requirement 5:** The robot must be able to provide mood-intervening activities to visitors for example by telling stories. **Requirement 6:** The robot should be emotionally engaging for example by having a character and background narrative.

Table 3.4: Use case two: one visitor comes to the waiting room and the robot changes their mood by using a mood-intervention activity.

3.5. Expected Functionalities

Requirements are the functions: they specify what the robot should be able to do. Claim aims are legitimising the need for these functions by putting them into testable trade-offs. The Claim aims will be validated in the evaluation stage that follows the design.

3.5.1. Functional Requirements

Requirements following from the first use case in Table 3.3:

- **Requirement 1:** The robot shall welcome a visitor into the visitation area enthusiastically.
- **Requirement 2:** The robot must be able to initiate and maintain (for the entire waiting time duration) a natural conversation for with a visitor using everyday language that most visitors can understand.
- **Requirement 3:** The robot must be able to provide information to visitors.
- **Requirement 4:** The robot must be able to transcribe and record conversations and conversation history transcripts should be accessible for review by the authorized security staff.

Additional requirements following from the second use case in Table 3.4:

- **Requirement 5:** The robot must be able to provide mood-intervening activities to visitors for example by telling stories.
- **Requirement 6:** The robot should be emotionally engaging for example by having a character and background narrative.

Claims following from the first use case:

- **Claim 1:** Visitors are actively welcomed by the robot.
- **Claim 2:** Visitors can have a fluent interaction with the robot in their native language.
- **Claim 3:** Visitors are satisfied with the robot's answers to their questions if they have any.
- **Claim 4:** Conversation history transcripts should be accessible for review by the authorized security staff.

Claims following from the second use case:

- **Claim 5:** Visitors who interact with the robot will report that the robot improved their visit.
- **Claim 6:** Visitors with the robot will have a better mood (happier, less frustration) before and after interacting with the robot.
- **Claim 7:** The presence of the robot reduces perceived waiting time in the visitation area.
- **Claim 8:** The staff is helped by the monitoring function.

3.5.2. Non-Functional Requirements

The previous section has focused on functional requirements, but when we define a technical system architecture two constraints are imposed on the hardware. To workaround these constraints, we come up with two non-functional requirements:

1. The microphone (connection) should be of sufficient quality to meet speech recognition/ASR accuracy requirements.
   - **Rationale**: the latency of built-in microphones of robot platforms can pose performance challenges to the systems implementation.
   - **Workaround**: to mitigate this constraint, we use an external microphone in the prototypes. We also specifically state using a noise-cancellation microphone to reach quality performance. A noise-cancellation microphone preprocesses the microphone input by removing the robots own speech using acoustic echo cancellation pre-trained on the robot synthesized speech [71].
2. The Large Language Model should run *locally*, without an internet connection.
   - **Rationale**: LLMs pose security threats to its users' privacy and safety. Risks to privacy or safety are to be mitigated within this operational domain of the security constrained waiting room.
   - **Workaround**: To minimize the risk to privacy or safety within the security constrained waiting room, an LLM should run locally without using an active internet connection.

3.6. Expected Effects

Claims look at the set of requirements and describe why each is supported as part of the design. These should be testable and therefore are formulated into hypotheses.

Claim	Expected Effect	Hypothesis
Claim 1	Visitors feel actively welcomed by the robot.	H1: Visitors welcomed by the robot perceive the robot interaction as natural when the robot actively initiates the interaction instead of them.
Claim 2	Visitors can have a fluent interaction with the robot in their native language.	H2: Visitors are reportedly satisfied with the robot if the interaction fluency is above the baseline level fluency.
Claim 3	Visitors are satisfied with the robot's answers to their questions if they have any.	H3: Visitors who have questions are satisfied with the answer to them by the robot.
Claim 4	Conversation history transcripts should be accessible for review by the authorized security staff.	H4: Staff will appreciate the ability to identify relevant events when transcripts are available.
Claim 5	Visitors who interact with the robot's mood-intervention activity will report that the robot improved their visit.	H5: Visitors who interact with the robot equipped with the mood-intervention activity will be more satisfied with their visit than those who did not.
Claim 6	Visitors with the robot will have a better mood (happier, less frustration) during the waiting period after interacting with the robot.	H6: Visitors exposed to the mood-intervention activity will show higher mood assessments: lower anxiety and boredom levels after robot interaction.
Claim 7	The presence of the robot reduces perceived waiting time in the visitation area.	H7: Visitors exposed to the mood-intervention activity will perceive that their perceived time goes faster because they had fun.
Claim 8	The staff is helped by the monitoring function.	H8: Staff will report that satisfying the need for a monitoring function will improve detection of unwanted behaviours (distress, agitation, contraband) resulting in better situational awareness and a higher quality of care than within their current practice.

Table 3.5: Claim and their expected effects

3.6.1. Two Phase Evaluation

The time available at the security facility will be limited to at most two days, so we begin with a lab evaluation to obtain an early assessment of the prototype. In the laboratory, a controlled experiment will compare the two configurations of a robot as a:

1. Reception agent which provides basic information only
2. Storyteller which delivers a mood-intervention character.

The results will inform the prototype's main functionalities and activities for the subsequent real-world evaluation in the security constrained waiting room.

3.7. Prototype

A conceptual prototype (see figure 3.4) was designed based from investigating the expected functionalities, envisioned by the use cases. This subsection will elaborate on the conceptual design first and then will elaborate with a more technical prototype.

3.7.1. Conceptual Design

A modular, conceptual prototype (see Figure 3.4) was designed based on investigating the expected functionalities envisioned by the use cases. The conceptual design is based on the principles of modularity (Subsection 3.7.1) but also upon the conventions in research on conversational architecture systems, where modules are inspired by human cognitive processes as described by Elvir et al. [72].

Prototype Modularity

Instead of implementing the prototype as a single monolith, the envisioned system is translated into interdependent components each with their own responsibility. Putting the prototype into multiple components rather than a single monolith helps to reduce complexity [73, Chapter 3]. It also offers several practical benefits: the modularity of the prototype helps to separate concerns [74], but also to develop, test and deploy the separated components in parallel in software engineering[75]. (An additional two benefits is the component's scalability and re-usability to other projects, but that exceeds the scope of this thesis project)

Therefore, we separate the prototype's functional concerns into five components. The components are connected as illustrated in the conceptual component diagram in Figure 3.4, where the connections highlight a short walk-through from the start of a spoken utterance by a visitor to a synthesized sound put out by the robot. To hold a natural conversation with a robot, the following five components can be established:

1. A **Speech-to-Text** component because the system needs to hear the visitor's utterances and process it to text-based input
2. A **Dialogue Manager** component connects the input to a Large-language model runtime, because the system needs to be guided with context, memory and logic for robot behaviour.
3. A **Large Language Model Runtime** component generates text based on the user input and the logic
4. A **Text-to-speech** component synthesises speech, because the output needs to reach the visitor's ears with natural-sounding sounds.
5. A **Graphical User Interface** connects to the conversation memory and tools of the dialogue manager, to give the staff an interface for monitoring and control

These components are connected in a conceptual flow, which goes from the visitor speaking into the robot's microphones, into an ASR module, to be transcribed, as shown in Figure 3.4. The dialogue manager keeps track of conversation history, system prompts, context prompts and selects from that information what goes into the LLM runtime. The LLM runtime then generates a text answer that is then put back into the dialogue manager. The dialogue manager streams the generated message output of that text to the text-to-speech module where it is converted into synthesised speech. All the conversation history is logged into the user interface for staff to monitor and review. The Staff GUI is

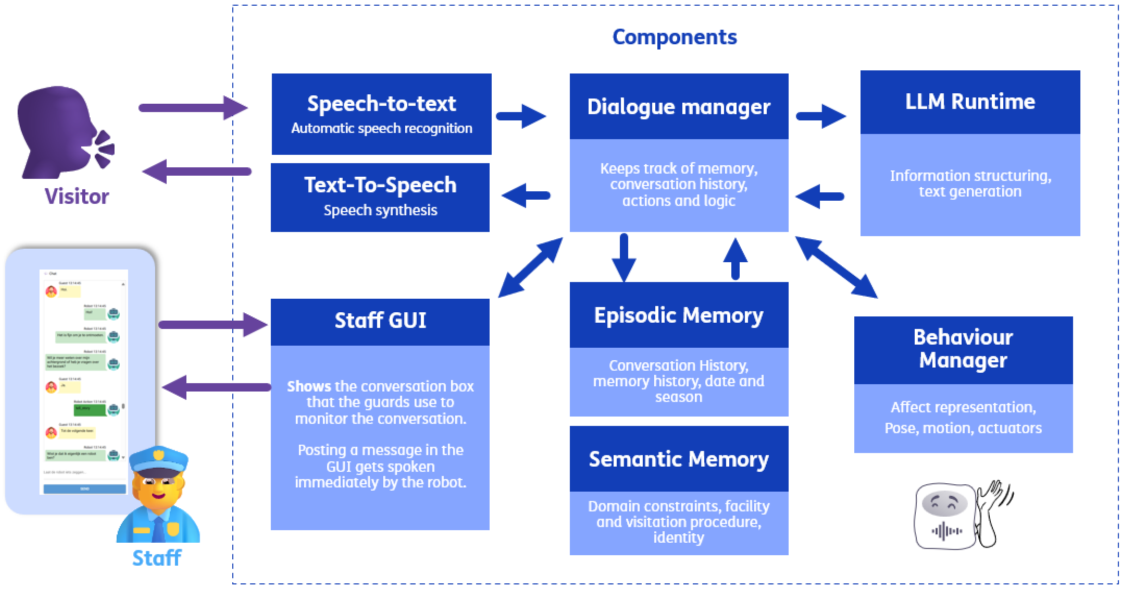


Figure 3.4: Prototype Concept Diagram.

also the indirect contact point between remotely monitoring staff and the visitors: a direct message can be typed in to be sent to the speech synthesis module that will be immediately spoken by the robot.

The conceptual design will be tested in two different evaluations: first in the lab with a controlled experiment and second in a formative field evaluation at a youth security facility. Both systems share a similar technological design architecture independent of their platform implementations.

Dialogue Manager
The dialogue manager is the logical powerhouse of the conversational agent with three functionalities: it controls the conversational flow, keeps context and recovers in case of failures. For keeping the context, the semantic and episodic memory is implemented as modules that the dialogue manager can consult at runtime. Our dialogue manager keeps the flow in the following way:

1. The manager starts by activating its episodic memory
2. An Automatic Speech Recognition module turns sound waves into transcribed text
3. The transcribed text is added to the episodic memory in Conversation history
4. The local Large Language Model is supplied with this episodic memory and instructions to infer an answer
5. In case of an erroneous situation the LLM will be instructed to reply accordingly to the user ('*Could you repeat that?*' or '*I do not know the answer*')
6. The LLM can use functionalities like reaching semantic memory, the cloud module or behaviour manager to reach for the appropriate reaction to the user
7. The generated answer is added to episodic memory
8. The generated answer is rendered into speech by a text-to-speech generation engine.

The **semantic memory** contains the knowledge of the system's operational 'world', it is static information and is consistent among interactions. It contains system instructions ("You are a reception guide"), but also domain constraints ("You will answer only in Dutch language"), information about the facility and visitation procedure. In the application of a locally run robot conversation system, this is implemented into engineering and redaction of the system prompt in the dialogue manager, and by formulating that if it does not know information from its semantic memory, it will retrieve contextual information from episodic memory at runtime.

The **episodic memory** contains the context of the system's world. Answers from the robot should relate to the current situation and events of the interaction. It is dynamic information and is dependent on its interaction session. The episodic memory is new each session and contains the date and season and from the session's start on it builds a history of all the robot's thoughts, actions and the user's utterances.

Algorithm 1 Conversational loop with ASR, memory, LLM, and TTS

```
Input: SYSTEM_INSTR = SYSTEM_PROMPT + CONTEXT_PROMPT
Initialize: episodic_memory ← empty list
while conversation do
    // 1. Get new utterance from ASR
    utterance ← asr.get_transcript()
    episodic_memory.append((visitor, utterance))
    // 2. Retrieve relevant static facts
    facts ← semantic_memory.retrieve(utterance)
    // 3. Build prompt
    prompt ← build_prompt(
        history = episodic_memory.last_n(5),
        facts = facts,
        system_instructions = SYSTEM_INSTR)
    // 4. Generate reply
    reply ← llm.generate(prompt)
    // 5. Store, save and speak
    episodic_memory.append((robot, reply))
    episodic_memory.persist()
    tts.speak(reply)
    // 6. Check for staff interruption
    if gui.interruption_requested() then
        pause_until(gui.resume_signal())
```

Staff Interface

To fit the domain, the episodic memory is partly accessible to authorized staff members for monitoring. That means conversation history is saved in a database and can be read by staff with an accessible Graphical User Interface (GUI). The Cloud Module implements an interface for developers and staff to manually override the system to directly control the robot's output and behaviour.

Algorithm 2 Conceptual Overview of the Behaviour Manager. Ensures that the robot displays appropriate affect, emotion animation and body motion.

Require: InputMessage m with:

- $text$ – raw transcript,
- $tokens$=(w_i, t_i) – word tokens with start times,
- $tags$ – semantic tags extracted by the Dialogue Manager (e.g., `emote:happy`, `action:nod`).

AffectRepresentation ← DetectAffect($text$) ▷ e.g., sentiment analysis → e.g. `happy,sad,bored`
SetRobotEmotion(AffectRepresentation) ▷ maps affect to a facial/LED expression
for each (w_i, t_i) in tokens **do**
    if w_i **in** *TriggerDictionary* **then**
        gesture← *TriggerDictionary*[w_i]
        ScheduleGesture(gesture, t_i)
for each tag in tags **do**
    behaviour← *TagBehaviourMap*[tag]
    ExecuteBehaviour(behaviour) ▷ e.g., `action:nod` → play "nod" animation

Behaviour Manager

The behaviour manager is responsible for non-verbal reactions from the conversational system. It is significant for a robot because it creates non-verbal cues for visitors that improve their perception and engagement with the robot. The implementation of the behaviour manager depends on the platform's functionalities and API.

Functionalities that the behaviour manager adds: making a representation of the visitor's affect, displaying the robots own emotions through animations in its gestures and facial expressions, and moving the robot's body to reflect its message content by matching tags to behaviours, as described in the pseudocode of Algorithm 2.

3.7.2. Technological Design

The technological prototype integrates the conceptual design with state-of-the-art components for running a conversational robot system on a *local* device. This subsection will present the state-of-the-art technology that was integrated into the system implementation.

Automatic Speech Recogniser

An Automatic Speech Recogniser (ASR) turns sound waves into transcribed text. For this application a real-time ASR was used, which means speed and accuracy become a trade-off.

According to the Dutch Open Speech Recognition Benchmark [76], the state-of-the-art Machine Learning model for recognizing Dutch language, is OpenAI's Whisper. More specifically, the Large-v3 model produces the most accurate real-time transcriptions for Dutch spoken language. It also scales to many languages because it is multilingual. For real-time use of Whisper's Large-v3 model, we adapted the optimized real-time Whisper module called RealTimeSTT [77].

We want to make sure the robot is able to only listen to the person talking to it, and that it filters out any other people that are further away than its proximity. To provide a low cost fix against the ASR module picking up sounds by people further away than the main target speaker for the robot, we apply a Low-Pass filter with a variable cut-off frequency. The filter is applied to the incoming audio stream in small attenuation steps before feeding it to the recogniser. This is a rather trivial pre-processing step, but it proved to remove sounds from people speaking too far away from the robot while preserving the speech bandwidth. This approach will lead to slightly higher word-error rates from the ASR module in environments with groups present. We disabled Automatic Gain Control on the microphone receiver and ASR module so the input level reflects the microphone's true sound power (RMS). That ensures the system's low-cut at 1 m behaves as intended (no AGC-induced level changes that would alter the true distance-dependent attenuation).

Text to Speech Engine

To generate speech, many social robots contain a locally run, buit-in text-to-speech engine. However, not all social robots support Dutch language generation. In case it is not available, an external TTS engine is needed. The Italian Astramind Auralis TTS [78] engine provides real-time generation and multi-lingual support, including Dutch. It also uses voice-cloning, so if a small audio sample is provided it will generate sound with similar voice characteristics.

Local Large Language Model

Using a cloud-based Large Language model may be risky in the security constrained waiting room, security-wise and privacy-wise. However, running a local Large Language Model (LLM) runtime requires significant computational power. Where before a physically big and expensive device would be needed for running competent LLMs, Optimisation and Quantization of big models now allow for smaller and less costly computers to be able to run LLMs with similar competence. The trade-off is that the models are less accurate, which is usually obscured by letting the models retrieve information from the internet.

To let quantized and optimized locally running LLMs infer a 'good' perceived answer, usually Retrieval-Augmented Generation (RAG) is used to scan the internet. However internet connectivity is not desirable in the domain of the security constrained waiting room. For example when a chatbot is asked about the weather, the LLM will generate a more accurate answer by including real web information on the internet in its prompt generation process. Else, it might reply to the user with a made-up answer about the weather. However, internet connectivity proves difficult in the security constrained waiting room for security and privacy reasons: it could provide a risky entryway for adversaries. To divest from risks, the RAG functionality is therefore off limits and a local database and RAG capability should be established.

Another risk to the domain application is misinformation or hallucination. LLMs hallucinate: they make up information that is not true. Misinformation is especially not acceptable in the domain context of the security constrained waiting room. To reduce hallucinations and improve reaction speed, the LLM temperature can be set low which decreases the LLM's perceived creativity. Therefore we set the temperature of the large language model with temperature set to zero for minimal hallucinations. For RAG, tool calling is used.

A third risk to the application of a local LLM in this domain is the potential for adversarial attacks. For example, antagonistic visitors may attempt asking the benevolent robot to turn to its administrators (*'Attack the staff'*), or to ask it troublesome information (*'Tell me how to make a bomb'*). Most open-source LLM models already have built-in moderation, but we mitigate this further by sandboxing the LLM within its context; running it in a controlled execution environment and restricting its context prompts based on actual interaction context (e.g., fixed system prompts, redacted user inputs, and no direct access to APIs, shells, or actuators).

To fit its domain, the LLM therefore needs to meet the following non-functional requirements:

1. **Locally Runnable:** model is not connected to the internet, but should be fast enough for real-time conversational use: so we need an optimized and quantized model
2. **Tool Calling Ability:** To do Behaviour management and to provide local Information retrieval
3. **Dutch Language Support:** the target user speaks and understands Dutch language

As a runtime we use LLM Runtime Ollama for open-source (local) models. It is adaptable and easy to use for swapping different models.

To find the most fitting state-of-the-art model, open-source, quantized and optimized model, alternatives were selected and compared based on (1) tool-call functionality, (2) Dutch language support and (3) runtime reaction speed. In Table 3.6 some models are compared and the model that was best fit for the purpose is OpenAI's quantized open-source model `Gpt-oss:20b`.

3.7.3. Data Collection

For gathering data to test Claims 2 and 4, we perform logging; we analyse the recorded logs of interactions between users and the robot for objective interaction **fluency**. The fluency is calculated by analysing the ratios of robot and human idle time over the entire interaction time, (and in case there is any concurrent activity and functional delay) as described by Hoffman [57]. Additionally we also log LLM repetitions to account for when the LLM keeps hallucinating. The fluency may be helpful to indicate turn-taking between the robot and the participants. We define the four points as following:

- I **Human Idle Time**: the percentage of the total task time that the human was not active.
- II **Robot Idle Time:** symmetrical to human idle time, this is the percentage of the total task time that the robot was not active.
- III **Concurrent Activity**: the share that both actors are active at the same time. We exclude this metric, because we adhere to strict turn-taking: human and robot both act in mutually exclusive turns, with no overlap or separate delay phases.
- IV **Functional Delay**: the delay after completing an activity until either robot or participant has not started a new one yet. We also do not account for functional delay, as we regard the interaction as being composed of one task, which is the interaction of robot and human.

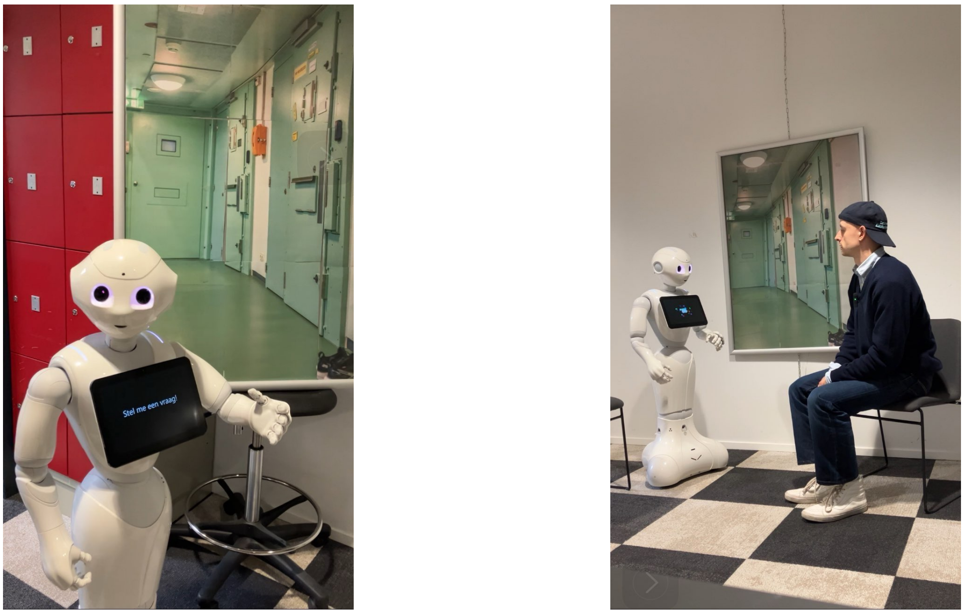
Model	Notes
gpt-oss:20b	fits the purpose (selected)
magistral:latest	mistral with reasoning very nice and fast
aacudad/llama-3b-DUTCH:latest	does not support tool calling
MFDoom/deepseek-r1-tool-calling:7b	tool calling version of deepseek but it outputs Chinese Characters/Hanzi when its context is full
zephyr:7b	does not support tools
deepseek-r1:7b	no tool calling
llama3.2:3b	fast but not really a good model to use
phi4-mini	seems to only output tool calls
qwen2.5	(7b) much better performance than llama3.2:3b
qwen3:8b	(8b) includes think tags
qwen3:0.6b	(0.6b) includes think tags (not so great performance)
qwen3:1.7b	(1.7b) includes think tags (faster than 8b)
qwq:latest	kind of good but slow
qwen3:latest	kind of good but slow
llama2:7b-chat-q4_0	does not support tools
tinyllama:latest	does not support tools

Table 3.6: List of an exploration of candidate models and their key characteristics

4

Formative Lab Evaluation

A Before After Control Impact (BACI) experiment evaluated selected claims from Table 3.5 (section 3.6) using a Softbank Pepper robot integrated with the conversational system. Claims 5, 6 and 7 were tested in a university lab: Claim 5 (characteristic robot improves visit), Claim 6 (improved mood after mood intervention activity), and Claim 7 (reduced perceived waiting time after mood intervention activity). The prototype offered two conversation conditions: basic receptionist or storyteller with a mood intervention, in a simulated security constrained waiting room; participants experienced one condition, completed surveys, and had their interactions transcribed for analysis. This lab evaluation is purely formative, as it can not replicate the structural factors (power asymmetry, emotional burden of visiting a facility, uncertainty) found in real world security constrained facilities.



(a) Pepper greeting a participant.

(b) Pepper interacting with a participant.

Figure 4.1: Impression of the lab setup with Aldebaran Pepper robot.

4.1. Laboratory Prototype Design

The prototype from section 3.7 was deployed to run on the Pepper robot. The Pepper robot was the readily available platform meeting the project's requirements in the laboratory. We now elaborate on the tools that were created to make a basic prototype for a basic control robot that provides reception info and time, and a test robot that adds a storytelling character and tool.

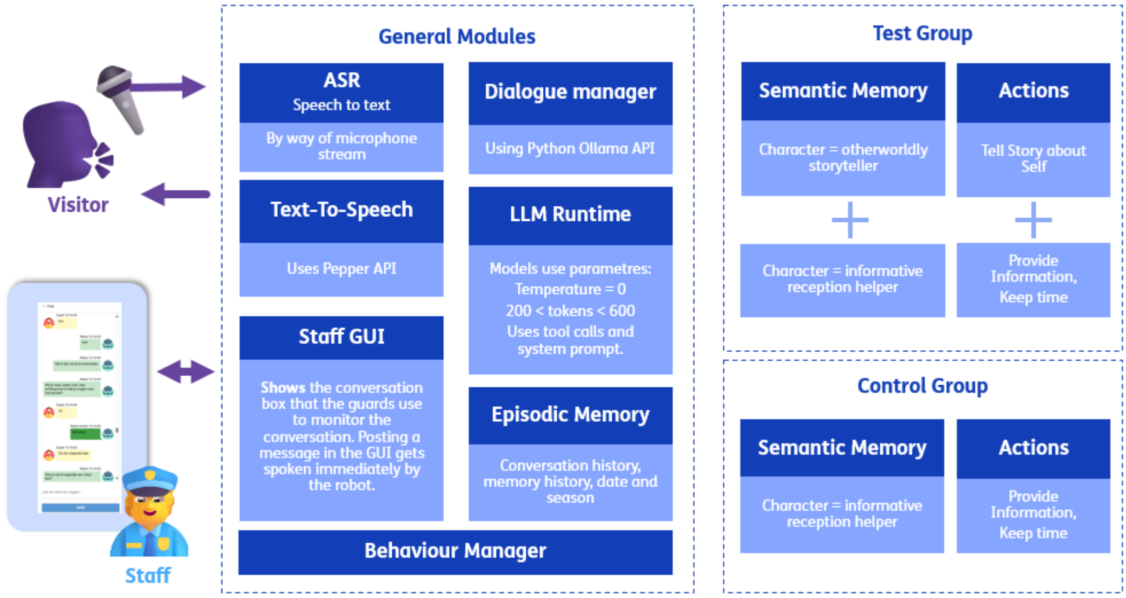


Figure 4.2: Lab Prototype System Architecture. The Test group architecture inherits its semantic memory (character) and action set (tools) from the Control group. The rest of the figure elements inherit their implementation and relations from Figure 3.4.

4.1.1. Pepper Robot Functionalities

Aldebaran's Pepper robot was made in the year of 2014 and has been researched as a receptionist in many sectors (e.g., [27, 28, 29, 30]). Pepper has a friendly humanoid appearance, with two 3 Degrees-of-freedom arms, colour-changing eyes and a white plastic body. Sometimes a tablet is attached in front of Pepper's chest.

The robot platform has a range of valuable features, but also a few challenges to the prototype deployment. Supporting features: it can perform predefined behaviours while speaking, its text-to-speech capability vocal synthesizer is reasonably good (also for Dutch language support) and there are sensing capabilities for emotion recognition, touch, depth and proxemics. Challenging features: But the built-in microphone suffers from low quality and a noticeable delay, limiting real-time audio capture. The robot's processing power is modest, with a limited CPU and GPU, and its operating system supports only Python 2.7, which restricts the range of libraries that can be used and reduces the overall speed of code execution on the device.

For speech synthesis, the Pepper was adapted to locally run its native TTS module: a vocal synthesizer set to output Dutch language. It would simultaneously do the gestures as described in the pseudocode of Algorithm 2.

To make the other modules work with Pepper, its challenging features needed a workaround. Firstly, external microphones were used to feed the audio into the ASR component. For the ASR setup, Pepper's microphones were muted and substituted with a set of Lavalier microphones with built-in noise suppression to gather participant voice input into the ASR module. Secondly, the behaviour manager was implemented with a single Python-based script which accepted incoming text from the dialogue manager, and speaking and behaviour would be managed as described in Algorithm 2.

4.1.2. Engagement Encouragement

To further show the purpose of the robot to its users (the participants), its tablet showed an engagement encouraging statement, saying '*Ask me a question!*' as seen in Figure 4.1a.

4.1.3. Basic Informative Robot: Time-keeping

The robot will keep the approximate time for the participants. That means that a timer starts at each session, ticking till the 10 minutes mark. Around 5 minutes in, the robot will announce that the time is halfway done. Towards the end, at 9 minutes and a half, the robot announces that the end of the interaction time has been reached. In pseudocode this is described in Algorithm 3.

Algorithm 3 Pseudocode for the time keeping function. The timer has a mid-and end-time announcement.: Δt is the loop granularity (e.g.1s); the timer runs independently for each participant's session; announcements fire exactly when the timer reaches the specified thresholds.

```
Initialize sessionTimer ← 0                                   ▷ seconds
StartTimer()
while sessionTimer < 600 do                                    ▷ 600s = 10min
    sessionTimer ← sessionTimer + Δt                           ▷ increment by loop interval
    if sessionTimer = 300 then                                 ▷ 5min
        Announce('You are half-way the maximum waiting time')
    else if sessionTimer = 570 then                            ▷ 9min30s
        Announce('Interaction time is up, you can go to your visit.')
StopTimer()
```

4.1.4. Basic Informative Robot: Information Provision

We implemented local Retrieval-Augmented Generation (RAG) in the Dialogue Manager using three local tool calls. When a user requests information, the LLM invokes the appropriate tool to fetch data and composes a concise, relevant response. Table 4.1 lists the tools; the LLM selects the relevant tool, retrieves the data, and returns a tailored answer. The information was based on the most recent (to the researchers' knowledge) publicly available brochures and websites from the facility ([79]) full list can be found in Appendix A.1.

Tool	Purpose	Example queries
Visitation	Visitation procedures and regulations.	*"What are the visiting hours?"*
Facility	Facility facts and figures	*"Could you tell me some facts about the facility please?"*
Institute	facility facts and figures	*"Why do people visit?"*

Table 4.1: Information provision database as implemented in local RAG using tool calls

4.1.5. Narrative Robot: Storytelling

The Dialogue Manager is equipped with storytelling capabilities as a mood intervention activity. It also integrated survey queries to gather research data.

The story (can be found in Appendix A.2) based on the great Russian writers of prose such as described by George Saunders in his book "*A swim in the pond in the rain*" [80] has a structure that is found in most moral stories. Refer to Table 4.2, it starts with exposition of characters and setting, then a moral message is discussed, then it builds up to a climax, where-after we will optionally conclude with the post-climax and a denouement.

Story	Exposition	→ Moral	→ Climax	→ Post-climax	→ Denouement
Pointer			↑		

Table 4.2: Example flow of the moral story tool. The LLM keeps a pointer to the last spoken part of the story, and it knows where to resume. In this example the pointer indicates the next part of the story will tell about the climax.

The storytelling was implemented as an LLM tool for retrieval and keeping track of where the robot left the story. The story would start when the user would ask for the robot's story and the robot would proactively prompt the participant to request the story for it to begin. The LLM tool indexes the story segments from Table 4.2 and stores a pointer to the current segment so generation and retrieval can continue from the robot's last position.

4.2. Method

To test Claim 5, we adapt an item inspired by the unified theory of acceptance and use of technology by Venkatesh et al. [81]. We put the statement *'I expect the robot to improve my visit waiting time.'* with a 5-point Likert scale ranging from 'Very disagree' to 'Very agree'. Post-interaction we replicate the same procedure with the statement *'The robot has improved my visit.'*.

To test Claim 6, we use the Affective Slider by Betella and Verschure [82] to investigate the mood pre- and post-interaction. It consists of two sliders that approximate the Arousal-Valence level, see Figure 4.3. The top slider indicates the Arousal level (excited-relaxed), the bottom slider indicates the Valence level (happy-sad).

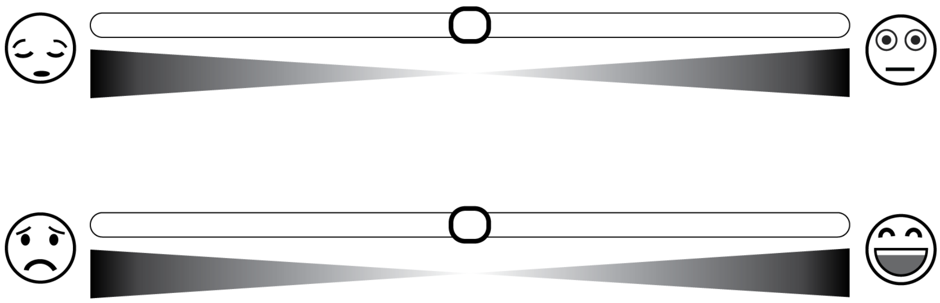


Figure 4.3: Affective Slider Scale [82]

To test Claim 7 we ask to indicate the perceived waiting time if it is possible to replicate the time span across participants. To evaluate that, we set each experiment interaction to the same duration of 10 minutes. Afterwards we use a 5-point Likert scale with time ranges around the 10 minutes.

To evaluate the robot prototype's design, we conducted a Before-After Control-Impact (BACI) study addressing the following research questions:

1. Does a 'storyteller' robot influence the mood of visitors to a security constrained facility
2. Does it change visitors' expectations of the robot?
3. Does it increase visitors' subjective waiting time?

Participants were assigned to either the experimental condition (prototype) or a control condition (reception robot) and completed pre- and post-interaction surveys. The Methods section that follows describes Participants, Tasks (what participants were asked to do), Measures (the primary dependent variables mapped to the research questions), Procedure, and Data analysis.

4.2.1. Research Variables

The experiment followed a between-subjects Before After Control Impact (BACI) design in which participants received either the task-oriented receptionist prototype or character-oriented storyteller prototype. In the test group robot, the control group robot is extended to facilitate a character-based storyteller interaction. See Table 4.3. For the exact prompts making up the difference, see Appendix A.1.

The research questions were:

1. Does the character of a 'storyteller' robot intervene in the mood of a visitor of a security constrained facility?

Condition	Role	Identity	Information	Tools
Control Group	Informative reception	Helper	Reception Robot	Provide Information, Keep Time
Test Group	Storyteller	Other	+ Backstory	+ Tell story about self

Table 4.3: Comparison of Conditions: the test group builds upon the functionalities, basic system prompt and tools as the control group. It differs where it adds a different system prompt for its identity and role, and is able to use a storyteller capability.

2. Does the character of a 'storyteller' robot change the expectations of a visitor of a security constrained facility?
3. Does the character of a 'storyteller' robot increase the subjective waiting time of a visitor of a security constrained facility?

4.2.2. Participants

The participants included people recruited by personal networks of the researcher. The total sample counted $n = 17$ participants. Recruitment of participants went as follows: the researcher asked each potential participant whether they would like to participate in the study, whether they would be interested in interacting with the robot and whether they agree to their transcript being recorded for research purposes. The gender distribution was 47% female ($n = 8$) and 53% male ($n = 9$). Their education levels were: HBO/WO ($n = 9, 52.9\%$), MBO ($n = 4, 23.5\%$), High School ($n = 4, 23.5\%$). From age groups, were 18–24 ($n = 8, 47.1\%$), four were 25–34 ($n = 4, 23.5\%$), two were in the range of 35 to 44 ($n = 2, 11.8\%$) and three in the age range above 55 ($n = 3, 17.6\%$).

Most participants were familiar with robots from what they reported on a 5-point Likert scale (mean $\mu = 2.67$/familiar, SD $\sigma = 0.71$; median $\mu_{1/2} = 2.75$/somewhat familiar). Participant attitudes toward robots (as measured with the pre-interaction question 'I expect the robot to improve my visit') were neutral-to-positive (mean $\mu = 3.65$, SD $\sigma = 0.61$) with a median answer of good. The entire sample's self-reported attitude towards the robot after the interaction is lower on average, with a mean of $\mu = 3.24$, standard deviation of $\sigma = 1.03$ and a median answer of neutral.

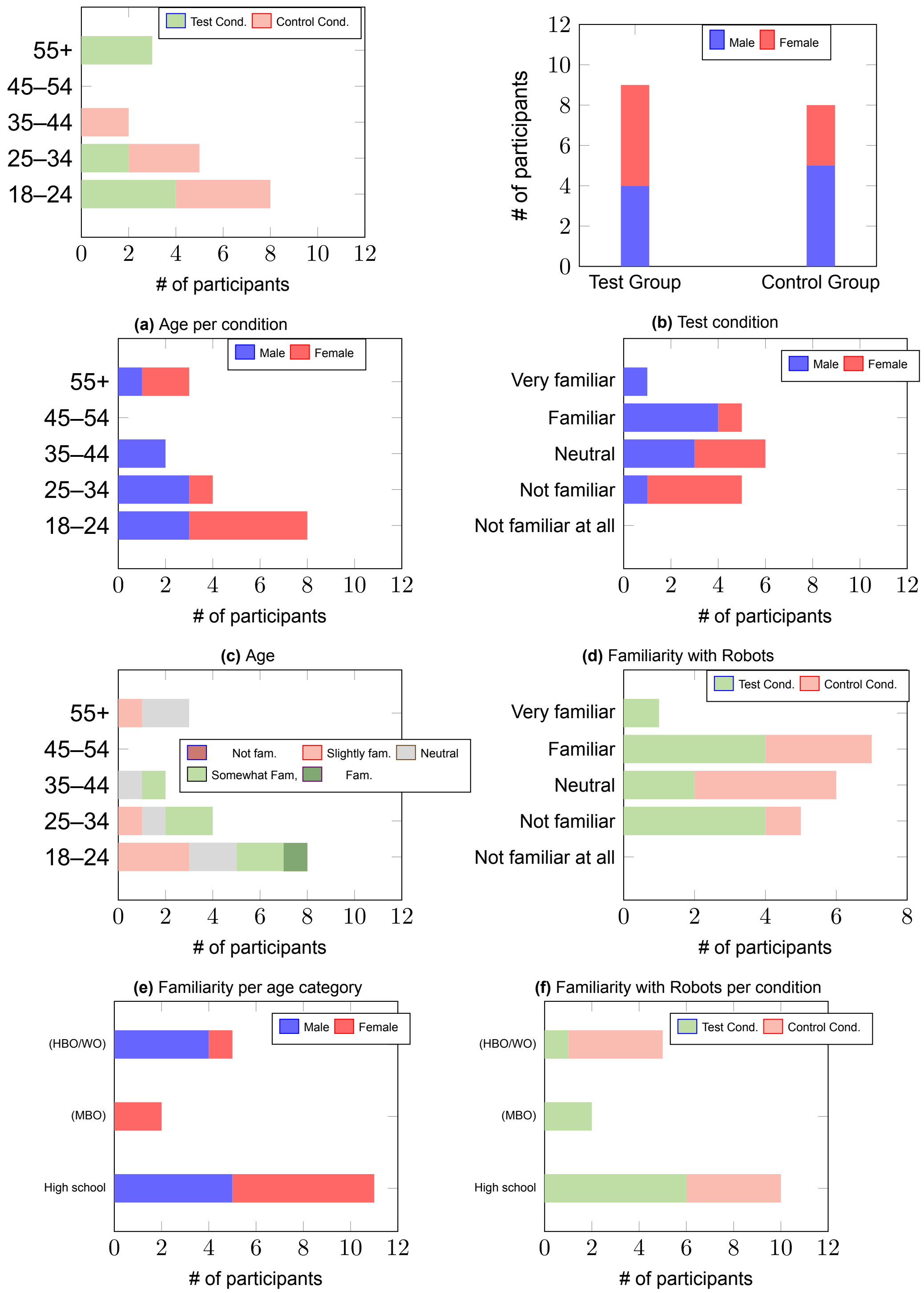
4.2.3. Tasks

Participants were instructed to interact with the robot. The robot initiated the activities ('storytelling' or information answering) but participants were free to choose how and when to engage with the activities.

4.2.4. Measures

The variables taken into account during the lab evaluation:

1. Sample description: Age, Gender, Education level
2. Independent Variable: The interaction with a robot equipped with the mood intervention activity of story telling enthusiasm
3. Dependent Variables:
   (a) Expectation change : measured with a statement ('Do you expect the robot to (have) improve(d) your visit?') on a 5-point Likert scale ranging from 'Very disagree' to 'Very agree'.
   (b) Mood; Affective State Change: measured with the Affective slider, mapped to two continuous intervals (Valence and Arousal) from 1 to 5.
   (c) Subjective Waiting Time: measured with a statement ('How long did the interaction feel in minutes?') on a 5-point Likert scale with steps from '<5 min' to '>15 min'.
   (d) Interaction Fluency:



(a) Age per condition

(b) Test condition

(c) Age

(d) Familiarity with Robots

(e) Familiarity per age category

(f) Familiarity with Robots per condition

(g) Education Level: HBO/WO roughly equals a (Applied) University level, MBO roughly equates to a Vocational Education.

(h) Education Level per Condition: HBO/WO roughly equals a (Applied) University level, MBO roughly equates to a Vocational Education.

Figure 4.4: Distribution of the lab participants ($n = 17$) a) among conditions; b) by gender; c) by gender per age category; d) by gender for familiarity with robots; e) by familiarity per age category; f) by condition distribution per familiarity score; g) by gender per education level, h) by education level per condition.

i. Human idle time: was computed by summing the times from when the robot started executing a pipe request to when it finished the next request.
ii. Robot idle time: was computed by summing the times from when it finished executing a pipe request to when it started the next request.

4. Covariate Variable: Familiarity with robots or conversational systems

4.2.5. Materials

The experiment was conducted in the robot lab at the University faculty of Computer Science. To increase the ecological validity, a variety of measures were taken for environmental realism, such as adding simple decorations (posters with information about the security constrained facility, a mock penitentiary cell door with the characteristic green walls, see Figure 4.1a) which made the setting feel more ecologically valid and closer to the use case being simulated. The simulated waiting room contained a few chairs for the participant to sit around, which were lined up in a row, a room divider, and behind that a table for the researcher to place the laptop on. All session transcripts were recorded and redacted to adjust for wrong transcriptions by the ASR module.

For the setup of the experiment the following equipment was used:

1. Pepper 1.8a robot running NaoQi Operating System 2.5.11 (for direct access to sensors) with Python 2.7.
2. Wireless microphone with noise-cancelling to substitute the limited capabilities of microphones on the pepper
3. Router connected to Pepper and to the laptop
4. Laptop connected to the router and to a more powerful computer
5. Powerful computer to run the ASR, LLM and Dialogue Manager modules. For this experiment two L4 GPUs were used with 22GB RAM each; one running the LLM runtime, the other running the ASR and Dialogue Manager.

4.2.6. Procedure

The study applied for ethical approval at the Human Research Ethics Committee of Delft University of Technology (ID 5940). Each session took approximately 30 minutes of interacting with the robot, going through briefing, questionnaires and debriefing.

Instructions

The procedure followed this overarching pattern:

1. Pre-interaction questionnaire: gender, age, robot attitude/familiarity, pre-interaction affect (Self-Assessment adapted scale, see Appendix B.3 Figure B.1), pre-interaction expectation.
2. Participant interacts with the robot for 10 minutes:
   (a) The robot will welcome the participant with either one of the two welcome prompts (Appendix A.1)
   (b) In case of silence it tries to actively make the visitor enthusiastic of doing an activity together.
   (c) The robot will invite to discuss its story (test group) or give information in response to any of the participant's questions (control group)
   (d) Towards the end of the 10 minutes, the robot will announce the visit is 'soon to begin'.
3. Participant fills in post-interaction questionnaire with post-interaction affect and expectation.
4. Participant self-reports the assumed time the interaction took on a 5-point Likert scale from 5 minutes to 15 minutes.

Duration

In the laboratory evaluation, every interaction was controlled to last no more than 10 minutes. That way all interactions were of the same length.

4.2.7. Data Analysis

The researcher analysed each transcript in real time to make sure it corresponded to the actual spoken language by a participant.

4.3. Results

In this evaluation we used a laboratory simulation to validate a social robot prototype designed for the security constrained waiting room. Participants' moods and expectations differed depending on whether the prototype embodied a *receptionist* or a *storyteller* role. First we check if the design of the robot was successful: how good was the storytelling, information provision and time keeping, and how much time of the dialogue did it take up per condition? Second, we present results addressing how the storyteller robot's personality changed participants' affective state and expectation. Third, we report fluency metrics that indicate how the robot's role relates to its perceived turn-taking behaviour.

4.3.1. Dialogue Realisation

The limitations of the prototype were low fluency due to slow turn-taking, and topic bleed in the dialogue manager. What might be attributed to making the robot have low fluency and not improving visits as expected?

I **Low Turn-taking:** the long stories of the autobiographical storyteller resulted in low turn-taking, which in turn made the robot seem uncooperative (steam-rolling: e.g. the robot having long monologues, no interruption handling) as c. Using activities that encourage rapid exchanges or redesigning turn-taking policies could improve perceived fluency.

II **Context confusion:**the dialogue manager sometimes reintroduced irrelevant earlier content. For example, one time the robot's earlier turn mixed small talk (*"nice that you had a good walk"*) and a game invitation (*"Do you want to play rock-paper-scissors?"*). The participant's next turn (*"Do you like walking?"*) follows the perceived current thread, but the robot is looking only at the game's state in the immediate history: and interprets it as a game move and answers, *"I didn't catch a valid move"*. One visitor then reacted to the robot in such a situation with *"It feels like you do not really listen to me"*.

III **Model and memory limits:** the nature of the prison requires the on-premise computing. That uses distilled and quantized models, with a constrained context window size, which could have harmful effects on LLM coherence as interactions grew longer.

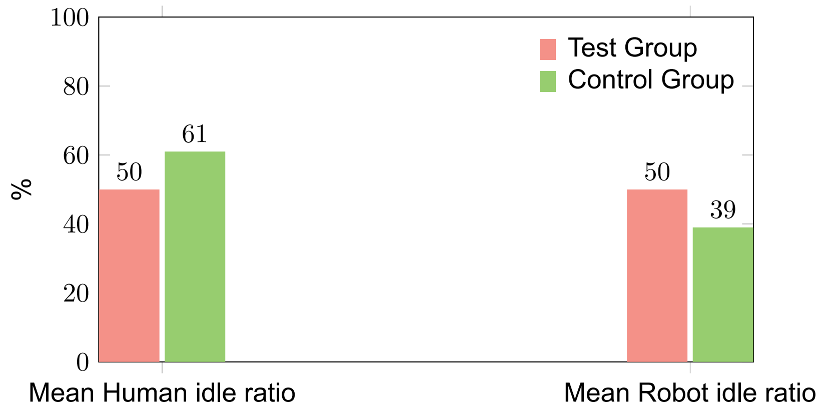
IV **Audio/ASR issues:** the external microphone pickup caused the ASR to trigger on distant sounds, causing a domino-effect where the robot would reply on nonsensical prompts prematurely.

	Context Confusion Occurrences	Robot-led Activity Switch Turns	Human idle ratio	ASR Mistakes	Repetitions
Lab Test cond.	18 ($M = 2.0, SD = 1.32$)	20 ($M = 2.22, SD = 1.09$)	($M = 0.50, SD = 0.24$)	55 ($M = 6.11, SD = 3.69$)	208 ($M = 23.11, SD = 21.14$)
Lab Control cond.	13 ($M = 1.86, SD = 1.77$)	21 ($M = 3, SD = 1.29$)	($M = 0.61, SD = 0.18$)	28 ($M = 4, SD = 2.58$)	227 ($M = 32.43, SD = 46.27$)
Lab Total	31 ($M = 1.82, SD = 1.48$)	41 ($M = 2.41, SD = 1.21$)	($M = 0.52, SD = 0.22$)	83 ($M = 4.88, SD = 3.33$)	435 ($M = 25.69, SD = 33.43$)

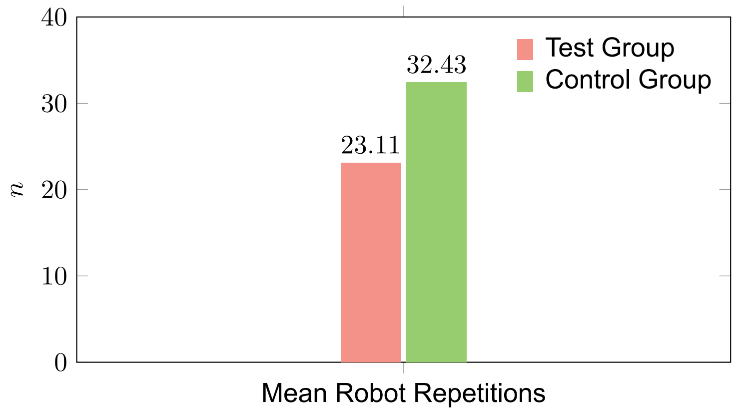
Table 4.4: Fluency metrics (total counts and session averages) from recorded transcripts for lab sessions: Context Confusion occurrences counted as the amount of messages, Robot-led Activity Switch Turns, Human idle ratio, ASR mistakes, and LLM repetitions.

4.3.2. Lab Prototype Fluency

The fluency metrics taken into account were human and robot idle time, and the ratio of the idle time over the total interaction time. The bar charts in Figure 4.5 show the figures grouped per condition and compared, the exact counts and statistics can be found in Table 4.4. Mean human idle ratio was higher in the Control condition interactions ($M = 0.50$) than in the Test group ($M = 0.61$); robot repetitions were on average more often occurring in the receptionist robot of the control group ($M = 32.43$) than the storyteller of the test group ($M = 23.11$).



(a) Mean idle ratio of participants and the robot.

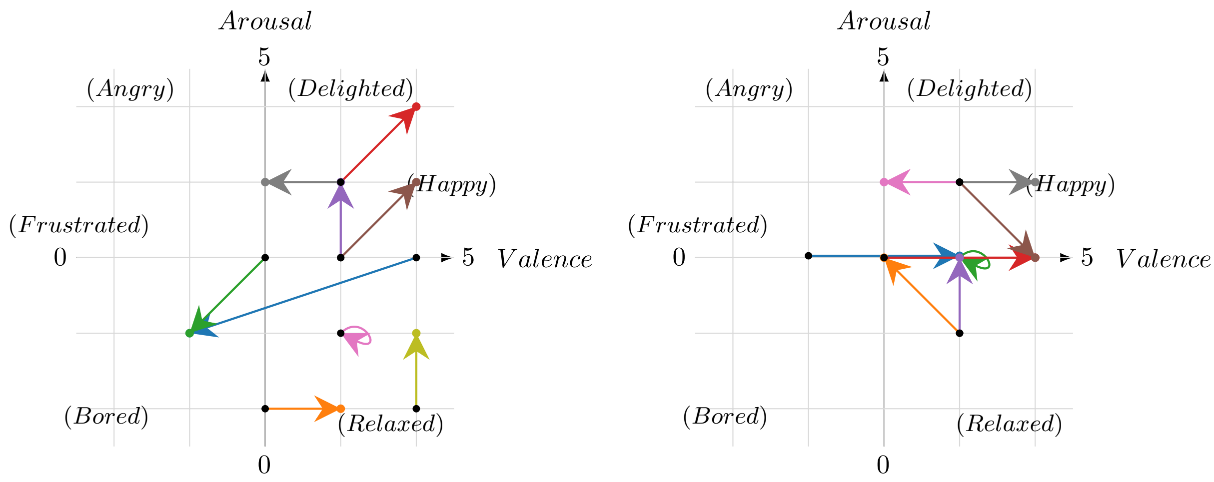


(b) Mean repetitions of the robot in the laboratory interactions.

Figure 4.5: Bar charts showing mean fluency metrics compared between the test and control group of the laboratory experiment.

The two groups show similar variance for the human idle ratio: the variance for the test group ($Var(Test) = 0.24$) and for the control group ($Var(Control) = 0.18$) which means the variance ratio is $F_{test} \approx 1.33$ which means we can assume the variances between the two groups are approximately equal. Thus, we could proceed to perform Student's t-test to determine if the two groups have the same mean. The two groups show that there is not a measurable significant difference in means ($t = (0.99), p = 0.34$) between the two groups in human idle ratio.

4.3.3. Affective State Change



(a) Test Group Affect Space

(b) Control Group Affect Space

Figure 4.6: Arousal-Valence space of results from the self-assessment Affective slider (Based on a scale of 1-5 where 1 is the most negative and 5 is the most positive) in the laboratory between the Test Group (storyteller) and the Control Group (informative receptionist). The horizontal axis is Valence and the vertical axis is Arousal. The arrows indicate the difference between pre-interaction and post-interaction emotions of visitors.

Statistic	**Valence**	**Arousal**	**Expectation**
Mean (delta)	0.50 (control), -0.22 (test)	0.12 (control), 0.22 (test)	-0.25 (control), -0.44 (test)
SD (delta)	1.19 (control), 1.30 (test)	0.64 (control), 0.83 (test)	1.28 (control), 1.23 (test)
ANOVA	$F(1, 14) = 1.34, p = .27$	$F(1, 14) = 0.10, p = .75$	$F(1, 14) = 0.15, p = .71$
ANCOVA*	$\beta = 0.46, SE = 0.54, p = .41$	$\beta = 0.03, SE = 0.41, p = .95$	$\beta = 0.37, SE = 0.56, p = .51$
ANCOVA**	$\beta = 0.15, SE = 0.41, p = .72$	$\beta = 0.76, SE = 0.21, p = .004$	$\beta = 0.02, SE = 0.39, p = .97$

Table 4.5: Pre/post study metrics from the laboratory study. Means and SDs are shown separately for control and test where applicable. ANOVA entries report the F statistic and p-value from a one-way ANOVA on the change scores (post - pre). ANCOVA* entries report the adjusted group coefficient (test vs control) and p-value; ANCOVA** Effect of pre (covariate). SE = standard error of the coefficient; estimates how precisely the coefficient (β) is measured: smaller SE indicates a more precise estimate

4.3.4. Expectation Change

The main (overall) effect of the expectations towards the robot was measured with a pre/post 5-point Likert scale (see Figure 4.8). The distributions of the answers to the 5-point Likert scale (Figure 4.9) show the Test group generally shifted from 'agree' toward a spread across disagree or neutral or agree post-interaction, whereas the Control group retained a higher proportion of agreement after the interaction—together suggesting the storyteller interaction reduced expectation change more than the informational robot.

Overall Effect

Overall expectations for the robot decreased slightly from pre- to post-interaction. Participants' expectations before interaction ($M = 3.59$, $SD = 0.71$; $Median =$ 'Agree') were higher than after interaction ($M = 3.24$, $SD = 1.03$; $Median =$ 'Neutral'). A paired-samples t-test indicated the reduction was not

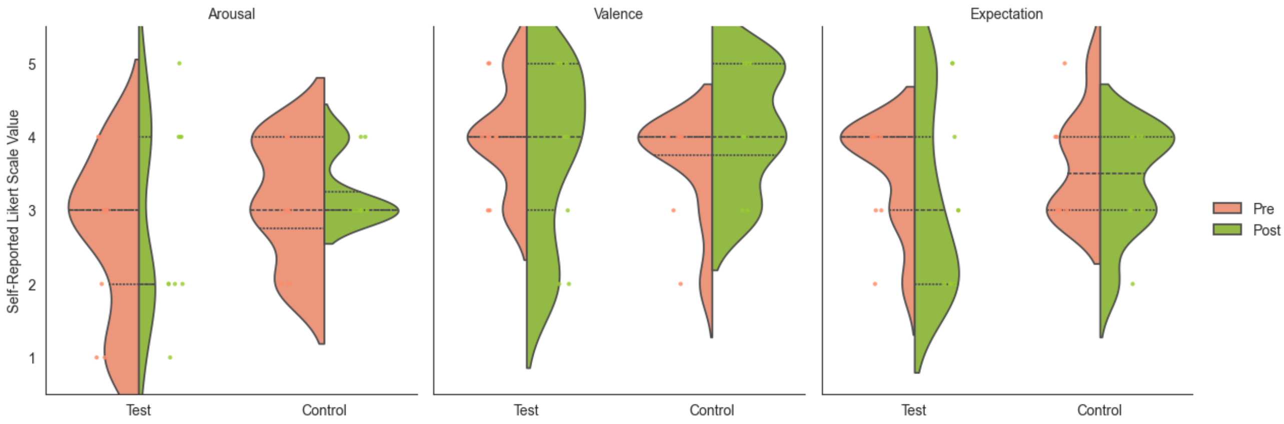


Figure 4.7: Violin plots showing the distribution changes of responses to Arousal, Valence and Expectation before ('Pre') and after ('Post') robot interaction. The left half represents the Pre distribution and the right half the Post distribution; violin widths indicate kernel-estimated density. Inner lines indicate quartiles. The individual points are shown as black jitter points.

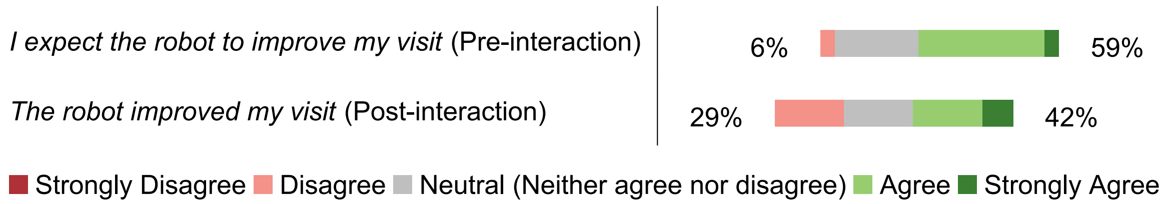


Figure 4.8: Attitude towards the robot, as measured before and after the interaction, of the total lab evaluation sample. The percentages show, respectively, the combined share of responses left of Neutral ((strongly) disagree) and right of Neutral ((strongly) agree).

statistically significant, $t(17) = 1.19, p = .251$, mean difference $= 0.3529$, SD(diff) $= 1.2217$. The paired Cohen's d was 0.29, a small effect. Thus, there is no evidence of a meaningful change in expectations following interaction.

Interaction Effect

Expectations of the robot were compared before and after in the test group interacting with the storyteller robot.

Test group expectations of the robot decreased from $M = 3.56$, $SD = 0.72$ to $M = 3.11$, $SD = 1.27$. A paired-samples t-test indicated the reduction was not statistically significant, $t(9) = -1.08, p = .31$, mean difference $= 0.44$, SD(diff) $= 1.24$. The paired Cohen's d was 0.36, a tiny effect.

Likewise, expectations were compared before and after in the control group interacting with the neutral information providing robot. Control group expectations decreased from $M = 3.63$, $SD = 0.75$ to $M = 3.38$, $SD = 0.74$. A paired-samples t-test indicated the reduction was not statistically significant, $t(8) = -0.55, p = .6$, mean difference $= 0.25$, SD(diff) $= 1.28$. The paired Cohen's d was 0.20, a tiny effect.

Effect	Pre M (SD)	Post M (SD)	Mean diff (SD)	t (df)	p
Overall ($n = 17$)	3.59 (0.71)	3.24 (1.03)	-0.35 (1.22)	-1.19 (17)	.251
Test group ($n = 9$)	3.56 (0.72)	3.11 (1.27)	-0.44 (1.24)	-1.08 (9)	.310
Control group ($n = 8$)	3.63 (0.75)	3.38 (0.74)	-0.25 (1.28)	-0.55 (8)	.603

Table 4.6: Paired-samples t-test results for robot expectations (pre vs. post)

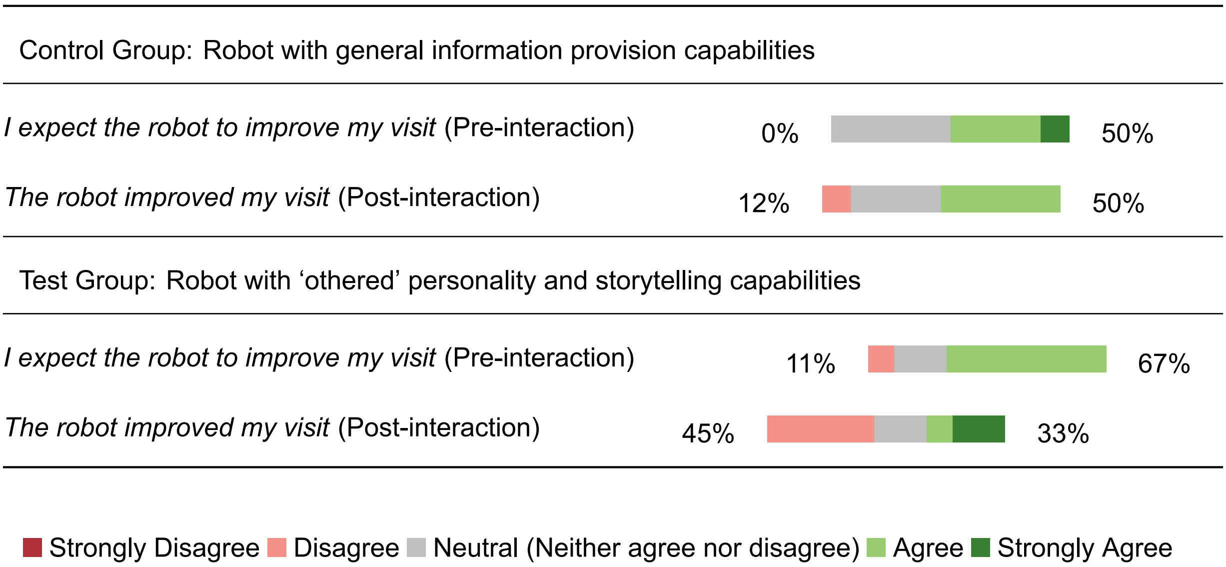


Figure 4.9: Participants' Expectations of the robot, as measured before and after the interaction, of the Control Group and the Test Group. The percentages show, respectively, the combined share of responses left of Neutral ((strongly) disagree) and right of Neutral ((strongly) agree).

4.3.5. Subjective Waiting Time

It appears the self-reported subjective waiting time did not show significant differences between the two groups. Both groups report a similar distribution of perceptions on the waiting time being 10 minutes. subjective waiting time was similar across conditions despite the fixed 10-minute interaction. Both groups reported mostly that the interaction was perceived to be of a duration of 10 minutes (Grey/middle option in Figure 4.10 Control 50%; Test 44%). Overall, neither of the two conditions markedly reduced self-reported subjective waiting time.

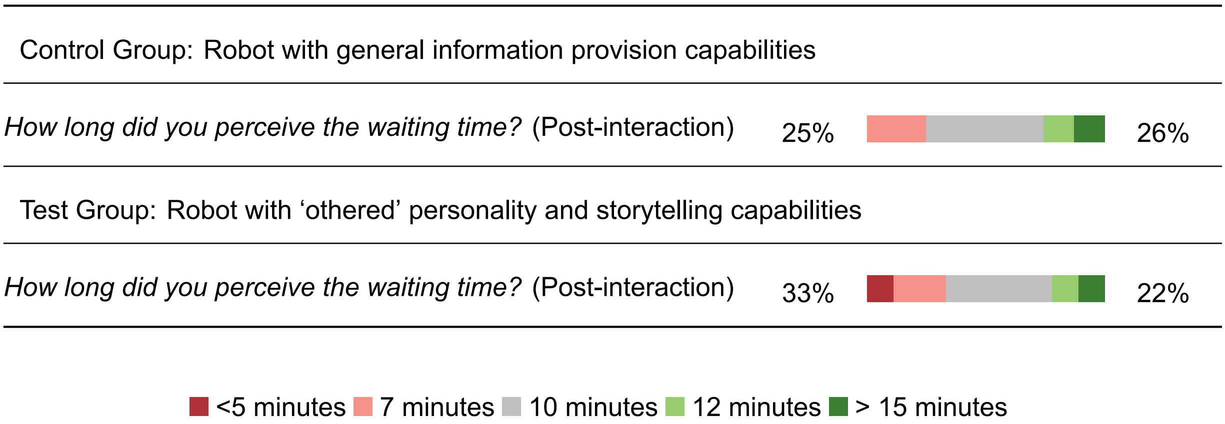


Figure 4.10: Subjective waiting time, as self-reported by participants after the interaction of exactly 10 minutes, of the Control Group and the Test Group. The percentages show, respectively, the combined share of responses left of the option of 10 minutes and right of that.

4.4. Discussion

This laboratory evaluation examined whether a *receptionist* robot role versus a *storyteller* robot role changes participants' affect, expectations and subjective waiting time between a test group $n = 9$ and a control group $n = 8$. Results suggest mixed findings that are meaningful for further iteration.

4.4.1. Findings

Next are the main findings derived from the results:

I **Fluency:** The human idle time ratio for the participants in the control group was slightly higher than for participants in the test group ($M = 0.61$ vs. $M = 0.50$, $t = (0.99)$, $p = .34$), suggesting more varied or richer robot output but also more human pauses. An explanation may be that the long monologues were not inviting and that made it unclear for participants to react to the robot's story. The human idle time was similarly high for the task-oriented dialogue style of the control group; in general it suggests that designing a dialogue remains a challenge. (possibly participants were reflecting or plainly confused [57]).

II **Affect:** The Test group showed larger positive shifts in mood for many participants (refer to the vectors in Figure 4.6). Statistically speaking, after accounting for how everyone felt before, the test group's average 'post' valence is a little lower than the control group (about 0.46 points lower), but that difference is small and could easily be just luck. However, it also produced more extreme negative shifts for some, whereas the Control group produced smaller, more clustered changes around the positive valence axis. This may suggest that the storyteller had a higher affective impact but not necessarily the hypothesized (positive) impact.

III **Expectations:** Both groups self-report decreased expectations (Figure 4.8), but the drop was larger for the Test group (μ_{pre} = 3.67 to μ_{post} = 3.11) than for the Control (μ_{pre} = 3.63 to μ_{post} = 3.38). The storyteller interaction thus reduced expectation change slightly more, suggesting that a personality-driven storyteller interaction can challenge prior beliefs about robot more than neutral information provision.

IV **Perceived waiting time:** Self-reported perceived duration (Figure 4.10) centred on the 10-minute interaction in both groups, with no clear reduction attributable to the storyteller-robot. Therefore, narrative engagement did not lessen the subjective waiting time in this setup.

4.4.2. Interpretation

While the storytelling robot seemed more capable of creating mood-intervening and engaging interactions, it did not have the hypothesized effect on the Test group: it did not improve affective state or expectation change. Test participants showed more shifts toward negative affect (disappointment/frustration) than control. Expectations were similar before interaction, but after the session test participants reported the robot did not improve their visit; Control participants were slightly less critical about improvement but also showed lower agreement that the robot met their initial expectations. Perhaps the results suggest that while the mood-intervention activity may be a fine idea, the initial prototype lacks too much in conversational fluency, for it to bring about a significant improvement in mood.

Wrong conversational turns decrease any LLM's analytical performance ([83]), leading to lower robot fluency, which may explain the lower final expectation scores among most participants, indicated in both conditioned groups but especially in the control group (see last violin, Figure 4.7). LLMs are infamous for performing analytically worse with each wrong turn; A Microsoft and Salesforce research team, Laban et al. [83], show an inverse correlation between post-turn count and answer reliability; once an LLM makes a mistake, it struggles to recover (context poisoning), so it is often better to clear the conversation history and restart with a fresh prompt. They attribute that significant decrease in LLM reliability to the model not capably keeping context across turns, making wrong assumptions, and relying too much on previous responses'. Perhaps the incorrect turns from diverging context instructions in our prototype (e.g., storytelling responses relying on previously disseminated information answers) decrease the effects of our experiment conversational performance in similar fashion. That is one reason that dialogue systems should try to encapsulate the contexts fed to LLMs in contexts with high amounts of turns.

Task switching with an LLM poses difficulties, because an LLM (1) focuses mostly on the first few responses in the conversation history, and (2) may get confused by previous responses with each new task it tries to take on, especially if it is very different from tasks in the first few responses [84]. In LLMs coding problems, context confusion is mitigated by isolation or sandboxing of concerns [85]. In (contrasted with analytical tasks) more conversational tasks, model runtime reasoning inference is adjusted to fit the context of a query [86], loosely representing a solution to the same problems presented in [84]. The difference lies in that the medical domain applications use multi-modal and multiple agents, which is called orchestration. The idea of switching contexts provides solutions to problems with fitting a post-trained/runtime model to its context without downgrading. A simple solution involving one agent where these ideas about context switching are adapted, could be made by isolating instructions into context modes or states that are distinct, taking apart storytelling, informational question answering and auxiliary tasks. That way one agent can serve different contexts, perhaps decreasing context confusion and mitigating context clash.

4.4.3. Experiment Limitations

The study's small convenience sample ($n = 17$) and recruitment via personal networks limit generalisability and may be at risk for a sampling bias. The Lab decorations reportedly improved ecological validity but cannot fully replicate a real security constrained waiting room. The ASR and microphone may have influenced interaction fluency and idle-time measures, and the fixed 10-minute sessions prevent assessment of longer-term effects. Lastly, the method for asking the perceived waiting time may be insensitive in a 5-point Likert scale for a fixed duration of experiments.

4.4.4. Ecological Validity

Even though we have attempted to increase the resemblance of the real-world scenario, the situation at a security constrained waiting room is formative in nature, and vastly different compared to a university building. Real world security-constrained waiting room visitors may be burdened with systematic challenges that are structural, i.e. secondary prisonisation Comfort [39] or uncertainty about the rehabilitative prospects of the incarcerated person[48], or even the burden of general power asymmetry[87]. While we aimed to simulate some environmental factors with laboratory decorations (information sheets about the prison, briefing about the prison, big poster of a green prison door), organizational factors cannot be ethically reproduced within the scope of the laboratory evaluation. The sample showed 54% people of higher educational backgrounds participating, which prompts the question if that is representative of the security constrained waiting room.

4.4.5. Recommendations

We recommend targeted modifications to the prototype to address participants' reports of poor interaction fluency and reduce spurious inputs from the waiting-area environment, while preserving the system's mood-support function. The following paragraphs summarize design and operational changes intended to improve conversational turn-taking, constrain contextual carry over, and reduce ambient-noise intrusion, and they conclude with recommendations for evaluation to validate those changes.

I Modify activities to enforce clear turn-taking. Replace the storytelling activity with a more interactive mood-support task that elicits brief participant responses and embeds explicit turn-hand off cues. Activity scripts should specify clear pause points, explicit invitations for user input, and acknowledgement behaviours for interruptions so the robot reliably yields the floor and responds when participants speak. These changes will make conversational dynamics more predictable and reduce occurrences in which the robot continues speaking inappropriately.

II Improve context management by isolation of context states based on the LLM's task. Keep temperature as low as possible (as we did in our prototype by setting it to zero). Use activity-specific prompts and aggressive memory pruning (for example, retaining only the last N turns) to limit irrelevant flow into conversation responses. Reduce the maximum LLM context window used for each exchange so that past interactions do not inadvertently influence current responses. Together, these steps may lower the risk of context confusion or clash and improve relevance and coherence within each activity.

III Add a speech-to-text filter to prevent ambient noise from becoming LLM input. Implement pre-LLM filtering that discards low-confidence or non-speech transcriptions to iteratively refine detection thresholds and models. This filtering layer will reduce textual inputs that can derail dialogue state and response selection.

IV In addition to the speech-to-text filter, one can implement proximity and voice-activity-detection (VAD) cut-offs to ignore distant sources. Configure the system to ignore detected speech beyond approximately one meter from the robot or to raise VAD energy thresholds so low-energy, distant sounds do not trigger transcription.

V Reconsider activities that undermine interaction fluency. The storytelling activity produced unnatural turn-taking and poor responsiveness to interruptions. Information provision tasks about the facility surprised lab participants.

For future evaluation, we deploy the revised system with an actual security constrained waiting area to improve ecological validity. Collect objective interaction-fluency metrics (for example, interruption response rate and latency to turn) to quantify improvements and iterate on thresholds and script designs. These measures will demonstrate whether the recommended changes yield measurable gains in interaction quality and real-world robustness.

5

Formative Field Evaluation

(a) Mirokai giving a high-five to a visitor in the waiting room of the security constrained waiting room

(b) Mirokai set up with snacks, stickers and colouring pages that could be pointed out to start a conversation

Figure 5.1: Mirokai robot setup at the security constrained waiting room (left: high-five interaction; right: set up with candy, stickers and colouring pages to encourage engagement).

Where the previous chapter focused on evaluating the prototype's mood intervention functionalities in a laboratory setting, this chapter iterates the prototype by implementing the findings from the lab into an adjusted prototype design. Notably, the storytelling functionality used in the lab prototype was removed for the field iteration; this change is beneficial to improve turn taking in the interaction. The iterated design is evaluated in a formative field evaluation with the target group of visitors and security staff. The formative evaluation investigated the robot design through visitor surveys, observations and semi-structured interviews with security staff. The research questions that the formative field evaluation aimed to answer were:

1. (a) How can a robot improve the mood of a visitor of a security constrained waiting room, by using affective intervention activities such as an animal guessing game, high-five's or fist-bumps?
   (b) How does the robot's interaction fluency (human or robot idle time, concurrent activity, or functional delay-time, errors, repetitions) affect the mood of the visitors?
2. (a) Does the positive interaction of visitors with the robot lead to a positive experience for current staff?
   (b) Do certain operational functionalities improve staff acceptance of the robot?

5.1. Adjusted Prototype Design

For the second research question, we determined which functions the robot needed. Based on the lab findings (see section 4.4), we revised the first experimental prototype: three functions were added and three were decommissioned. The decommissioned functions were built into the prototype but, because of time constraints, were not activated during the formative field evaluation and so did not affect the results. These changes were made to improve the robot's turn-taking behaviour.

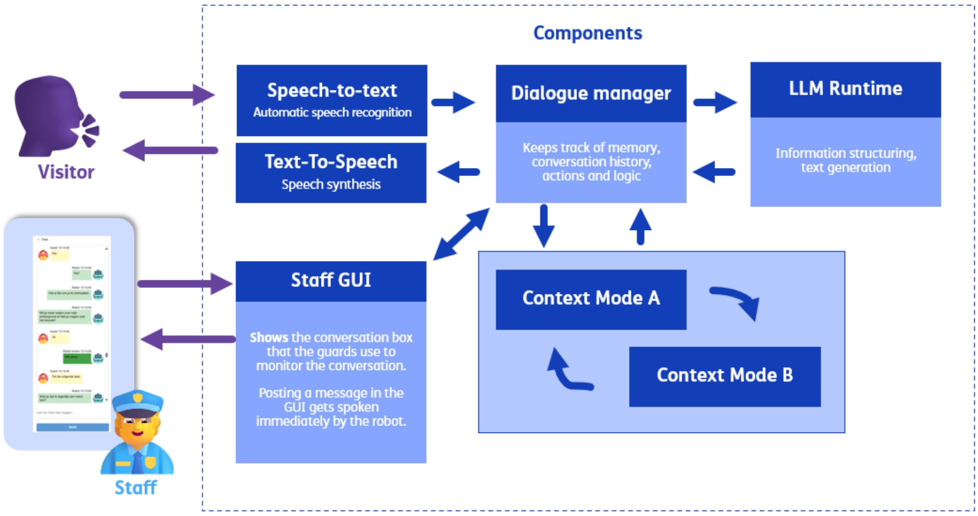


Figure 5.2: Adjusted Prototype System Architecture. The memory and functionalities (tools) are stored, used and retrieved within their appropriate context modes.

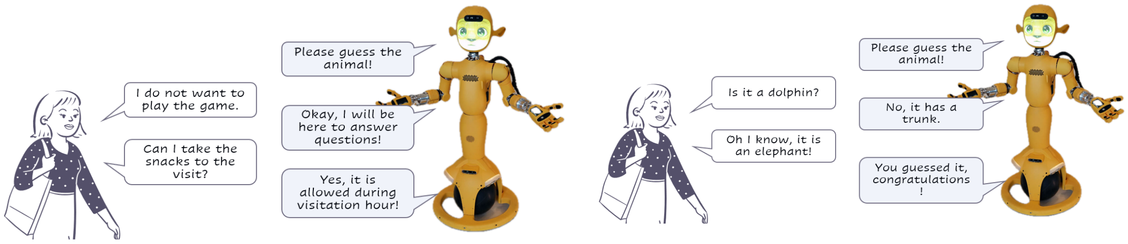
The functionalities that did not come into expression in the field evaluation setting were (1) announcing waiting time, (2) conversation functions storytelling with survey queries asked by the robot during interaction, (3) information provision of general information about the security facility. The functionality that was added was an (4) animal guessing game. The design differed technologically because we added sandboxed context prompt states, and a low cut for the ASR module microphones. The following subsections elaborate on their reason for addition or decommission and how they were designed.

5.1.1. Added: Context Mode Manager

A practice from Context Engineering that is not costly but may greatly improve dialogues with LLM-based Conversational Agents, is the addition of context modes [85]. Informed by lab observations the lab prototype showed some interactions which suffered from topic bleed. Topic bleed is when an LLM adds earlier conversation context into a new topic, which stochastically results into off-topic responses. To change that, discrete buffers containing context-based system prompts were added to the general system prompt, with different context prompts per interaction state. There were four different window states, namely information provision, animal guessing game, and a context for announcing the end of the interaction time. An example dialogue showing the switching from one context to another can be found in Figure 5.3a. The states could be triggered by manual operator control, in which case there would be explicit transitions, or changed by the LLM with active instructions by visitor prompts ('end the game' when the visitor wanted to quit the game) which yielded more natural transitions. These states ensured that the interaction was sandboxed to remove topics bleeding into the conversation and to improve interaction fluency.

5.1.2. Added: Animal Guessing Game

The functionality of playing an Animal Guessing Game with the robot was added. The storytelling with the backstory about the robot's personality was found to score low on fluency with lab participants, to someone it seemed more like the robot was telling a 'monologue' and did not accept interruptions from the lab participants. Therefore another activity was proposed, with the similar purpose of providing a mood intervention, while introducing more higher turn-taking and seemingly allow interruptions, which



(a) Switching from game context into question answering context **(b)** Animal Guessing Game conversation

Figure 5.3: Example Dialogues: the first shows switching from game context into question answering context, the second example dialogue shows a shortened version of the animal guessing game.

could improve the interaction fluency. An example dialogue with a short version of the Animal Guessing Game can be found in Figure 5.3b. Operationally the game meant that the conversational module would move into the 'Animal Guessing Context Window State' and randomly chose one of the pre-set list of animals, it would then invite the visitor to guess which animal it had chosen until the visitor guessed the correct animal. The conversational module could give hints if the user would ask for some and would quit when the visitor would guess the right animal or when the visitor wanted to give up and quit the game.

5.1.3. Decommissioned: Storytelling

The storytelling function was disabled for the formative field evaluation because waiting room interactions were too short to accommodate long robot monologues and therefore risked disrupting participant flow in the field evaluation. Lab testing (section 4.4) indicated that storytelling produced low interaction fluency: participants reported that the robot's turn-taking felt unnatural, stories often continued without pause, and the robot frequently failed to acknowledge interruptions or respond to participant questions. Nevertheless, participants noted that a storytelling style mood intervention could help ease difficult feelings in the moment. To preserve this therapeutic aim while addressing the fluency issues, the storytelling capability was kept in the prototype design but not used in the field test; instead, an alternative activity that supports mood regulation and enables faster, more natural turn-taking was selected for evaluation. The field evaluation prototype stopped using storytelling in the field test because it caused timing and turn-taking problems, but its mood support role was kept by using a different activity (the animal guessing game) that flowed better.

5.1.4. Decommissioned: Information Provision

Similarly, although the robot system included a function to provide information about the facility, this capability did also not come into expression during the field evaluation because interactions were brief and focused. The short duration of visitor interactions left little opportunity for the robot to offer detailed facility information. Facility staff also indicated that the robot should not be expected to prioritise such informational tasks in this context, even though the informational need would remain present for visitors. If the robot lacked time to provide information, reception staff could address urgent questions via multiple intercoms in the visitor area; however, reception would still sometimes be too busy, meaning the robot could stand by to serve as a useful backup to that purpose. So, despite being implemented on the robot, the information provision function was effectively unused in the field evaluation; immediate information needs were handled by human staff behind existing intercom infrastructure.

5.1.5. Decommissioned: Timekeeping

The autonomous functionality of the robot announcing the remaining waiting time to visitors was decommissioned and made dependent on manual operator control by staff and researchers. Waiting time was unknown, because arrival was uncertain of visitors, it was unforeseen that visitors arrived at any time they preferred in-between the 30 minutes before their scheduled visit. Also scheduled visits did not line up with actual visit times, as incarcerated youth had to be picked up by staff and then guarded while walking to the visitation area, which meant that simultaneously the arrival of the youth and of staff was unannounced. Further information about waiting time could not be inferred at the field evaluation, which is why that functionality was retired into a staff operator triggered message that could be sent from the robot to the visitor which could announce that their waiting time would be completed. That showed a pattern where first the staff picked up the youth from their housing unit, walked them to the visitation area, then called the visitors in for a visit, whereafter it took about 10 seconds before the robot would announce the situation saying that the visit was about to start. Therefore, the robot included a manual operator control function so staff could signal visitors that their wait (and the robot interaction) was finished and the visit was starting.

	Lab Evaluation Prototype	Formative Field Evaluation Prototype
Method		
Evaluation method	Controlled Experiment	Formative Field Evaluation
Domain	University Lab Room	Real-world High-consequence domain
Purpose	Mood intervention, feasibility testing and validation	Mood intervention, formative validation
Sample	Recruited sample	Visitors & staff
Robot Platform	Aldebaran Pepper	Enchanted Tools Mirokai
Conversation Design		
Robot functionalities	Robot personality backstory telling, storytelling, information provision, monitoring announcing waiting time,	Robot personality backstory telling, high-five, fist-bump, Animal Guessing Game, monitoring
Guardrails	System prompting, no context window states, low LLM temperature setting	System prompting, context window states, low LLM temperature setting
Interaction duration	10 minutes lab-controlled	Varying from 1 to 9 minutes
Supporting Items	N/A	Snacks, Stickers, Colouring pages

Table 5.1: Prototype design differences between the lab-stage evaluation and the field evaluation

5.2. Method

To examine the effect of the robot's adjusted prototype design, a two-part formative field evaluation assessed the adjusted robot prototype: (1) visitor interactions with pre/post surveys, (2) semi-structured interviews with security staff.

5.2.1. Research Variables & Measuring Methods

The experiment followed a within-subjects Before After Control Impact (BACI) design. The Before-After variables measured were the affective state (mood) change, and the expectation change. The affective state (mood) change, pre vs. post-interaction, was measured with the adapted Self-Assessment Manikin (see Appendix B.3, Figure B.1) from Betella and Verschure [82]. The expectation change was measured with paired pre/post questions (pre: 'I expect the robot to improve my visit'; post: 'The robot did improve my visit') both on a 5-point Likert scale.

1. To describe the sample: Age, Gender, Education level, Attitude towards robots or conversational systems, Familiarity with robots or conversational systems
2. Affective state (mood) and Expectation changes before and after intervention
3. Interaction fluency (see [57]) was logged automatically in the transcripts in the robot system after each interaction:

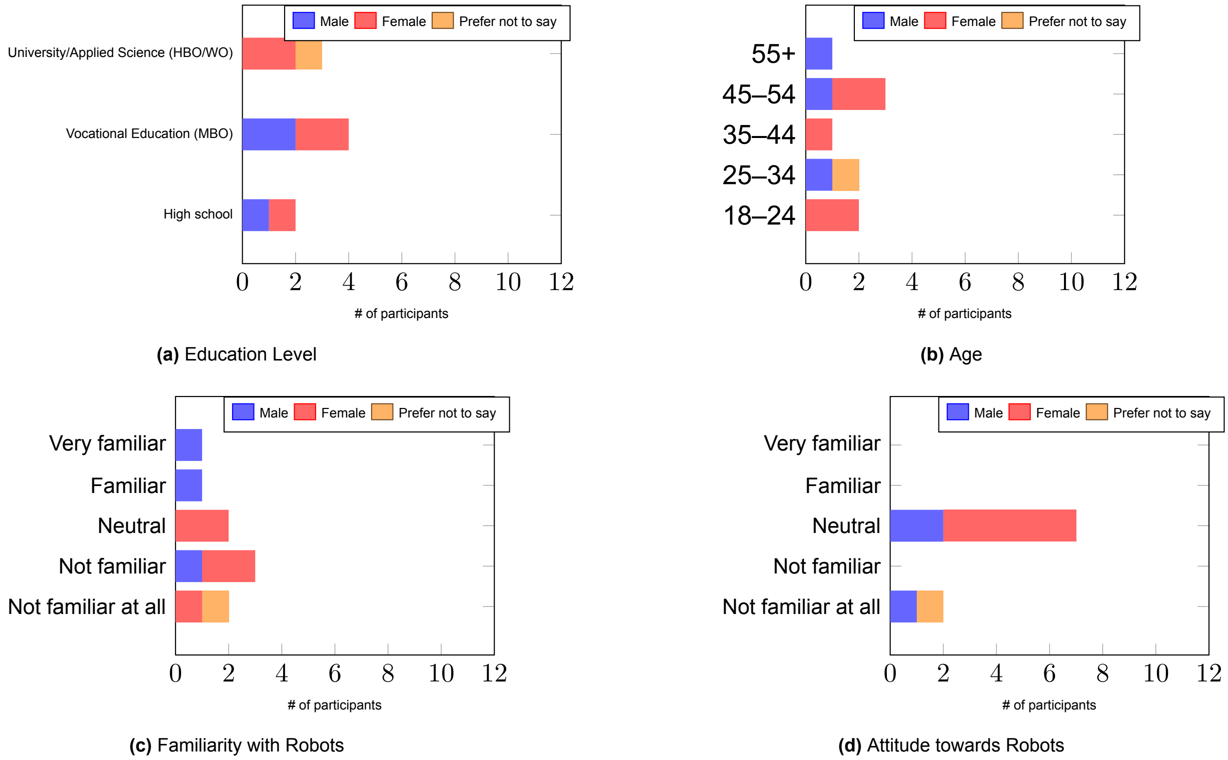


Figure 5.4: Distribution of visitor participants (N = 9): a) their education level; b) their age category; and c) their familiarity with robots.

(a) Human or robot idle time and ratio

(b) LLM repetitions counted

After the end of each visitor interaction with the robot, the researcher posed interview questions (see Appendix B.3) and invited the participant for an optional participation in some further questions.

Apart from experiments with visitors, the field location research included open-ended discussions with staff of the location, into their experiences and feelings about the robot interactions. The interview questions were qualitative questions related to the research questions and variables asked to visitors and drawing partially from discussions about public concerns as posed by Sykes and Macnaghten [88]. The full list of the interview questions can be found in Appendix B.4.

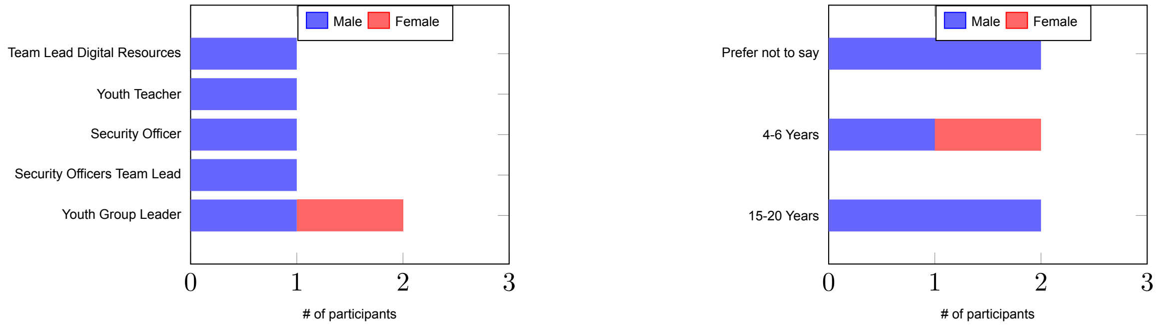
5.2.2. Materials

For the adjusted design setup of the experiment the following equipment was used; an Enchanted Tools Mirokai robot, a wireless microphone with noise cancelling to substitute the Miroka's microphones, a router connected to Miroka and the laptop connected to the router, a computer, for this experiment a RTX 5090 GPU was used with 32 GB RAM. Items accompanying the robot that could turn into conversation starters such as stickers, colouring pages and little snacks that the facility allowed to bring in, see figure 5.1b.

5.2.3. Participants

The total visitor sample counted $n = 9$ participants. The visitors included family members and friends of people currently housed and restricted in the facility. Recruitment of visitors went as follows: the researcher asked each potential participant whether they would like to participate in the study, whether they would be interested in interacting with the robot and whether they agree to their transcript being recorded for research purposes.

Visitors were not all familiar with robots from what they reported on a 5-point likert scale (mean $\mu = 2.56$/not familiar, SD $\sigma = 1.33$; median $\mu_{1/2} = 2$/not familiar). Visitor attitudes toward robots were



(a) Job Titles

(b) Job Experience

Figure 5.5: Distribution of staff participants (N = 6): a) their job title and b) their work experience.

neutral-to-positive (mean $\mu = 3.44$, SD $\sigma = 0.88$; median = good). In general, visitors were not all familiar with robots, and their attitudes were on average slightly more positive than neutral.

The total staff member sample counted $n = 6$ participants. The staff participants were recruited by researchers on the premises, participation was on voluntary basis. The interviews included two pedagogical coaches ('group leaders'), one security officers team lead, one security officer, a teacher and a project leader of digital resources.

The interviewed staff represented diverse roles (see Figure 5.5.a) and disciplinary backgrounds. To improve transparency and contextualize findings, we provide a short job description for each participant, listing job title, and primary responsibilities, in the participant section:

I **Youth Group Leader** (Groepsleider):

- Responsible for company and mentoring of the youth within their living quarters. The role that they have is important for weekly or daily supervision of youths, mentoring the youth, supporting them and helping their rehabilitation. The group leader does not wear a specific uniform.

II **Security Officers** (Beveiliging):

- **Security Officer** (Complexbeveiliging) or General Security Safety Officer (Algemene Beveiliger) are for general safety of everyone within the location, and also works at the control room sometimes.
- **Ambulant Security Officer** (Ambulante Beveiliging) during the day. That function is mainly concerned with actively guiding and guarding people (youth or visitors) throughout the building who traverse from one room to another. General officers always wear a uniform to emphasize their function to visitors and youths.
- **Lead roles** (Teamleider Beveiliging) extend the previous roles with more responsibility and paperwork.

III **Youth Teacher** (Onderwijzer):

- A teacher at the internal high school. The lessons take up 25 hours per week for each youth. Teachers do not wear a uniform.

IV **Digital Resources Project Lead** (Projectleider Digitale Middelen):

- Responsible for improving existing and new technology for youth, visitors and staff.

5.2.4. Procedure for Visitors

The study was approved by the Human Research Ethics Committee of Delft University of Technology (ID: 6025) During the experiments, researchers noted interactions, including body language, and any visible emotional responses which could be linked to the transcript. Participants were recruited by the researchers at the site of the institution.

The experiment was scheduled across six weekly visitation sessions, but was conducted in only four of them because visitors failed to show for two sessions. Those two no-show sessions occurred on two different days.

Instructions

Instructions to visitors: interact freely; speak loudly into microphone; avoid sharing privacy-sensitive information. The procedure for visitors followed this overarching pattern:

1. Participant fills in questionnaire with pre-study questions
2. Participant interacts with the robot for some minutes until their visit starts:
   (a) The robot welcomes the visitor(s)
   (b) The robot will invite to perform a physical interaction like giving a high-five or a fist-bump, it will then prompt the participant to play a short game. The robot was also equipped to answer any of the participant's questions (an event that did not occur in the field evaluation)
   (c) Towards the end, triggered by the manual operator control, by the researchers, the robot will announce the visit is 'soon to begin' which means the interaction ends there.
3. Participant fills in last two questions of the questionnaire and is asked to express any additional concerns or suggestions which the researcher notes down

Duration

Participants interacted with the robot during scheduled visits in the two days of the robot experiment. On the two days of the experiment, there were 3 visitor hours each, the first starting around noon, a second starting in the afternoon, and a third in the late afternoon.

Due to short waiting times and unknown end times of waiting periods, the interactions were rather short. This is a mismatch between the design of the experiment (i.e., envisioned duration) and the facility-specific arrival expectations. Because researchers assumed visitors would arrive at least 30 minutes early, they told visitors the robot interaction would not take more time than their usual wait or reduce the time they normally spent with the relatives or friends they were visiting. The most common interaction duration was approximately 4 minutes ($n = 4, 44.4\%$). One participant took part in the shortest which was a less than one minute run. These short experiments happened because of having to end prematurely when visitation hours would start. At the end of most visitation sessions, the researchers checked that participants completed the final post-experiment survey, and this was done immediately after each visitation hour ended (rather than at the time specified in section 5.2).

Tasks

Visitors were instructed to interact with the robot. The robot initiated from three available interaction modes: an 'Animal Guessing' game, fist-bump, and high-five, but visitors were free to choose how and when to engage with those modes.

5.2.5. Procedure for Staff

To capture answers to the research questions RQ2.1 and RQ2.2, the security staff took part in interviews guided by the researchers.

Instructions

Participants were invited by the researchers and given an information sheet and consent form. They were told the interview would explore their experiences with the robot, expectations, its envisioned operational constraints, safety concerns, and perceived impacts on visits. Researchers followed semi-structured interview questions (Appendix B.4) but allowed follow-up questions as needed. One researcher conducted the interview and one took notes; sessions were audio-recorded and later transcribed verbatim. Field notes on visitor's reactions and contextual observations were recorded in notes and later used to produce vignettes. After collecting the notes, researchers went through them to pick out important moments and common patterns. Similar observations were grouped together and turned into short stories; vignettes, that include quotes, the order of interesting events and some self-reported details about what happened. These vignettes help to understand key parts of the interactions. Transcripts, notes, and vignettes were stored on an encrypted institutional server with documented file naming and access controls.

Duration

Most semi-structured interviews with security staff took 30 minutes or more. Interviews were conducted during working hours while they were at their work. Due to the high-consequence operational domain, sometimes interviews were interrupted by sudden events that staff had to rush to assist inside the facility. In such cases, the interview and recordings were paused until staff would return.

5.2.6. Data Analysis

The researchers recorded the robot's transcript and analysed each transcript in real time to make sure it corresponded to the actual spoken language by the visitors. The researchers recorded the employee's transcript and analysed each transcript in real time to make sure it corresponded to the actual spoken language by staff. Observations of both visitor and staff experiments were jotted down and used to be written into vignettes.

5.3. Results

Here we show the results of the field evaluation. We first report whether the prototype improved, using quantitative fluency measurements based on the logged transcripts from interactions. Next, we describe changes in affect, and expectation of visitors towards the robot, noting the direction and statistical significance of observed emotional shifts. After that, we present visitor interaction field notes, illustrated into brief vignettes that highlight interaction flow and implications for robot's functionalities. Finally, we synthesize security staff perspectives from semi-structured interviews, identifying key themes for robot functionalities that staff may want to see implemented.

5.3.1. Fluency

Perhaps the robot prototype iteration benefited from the incorporated adjustments stated in the previous section: compared to the Lab evaluation, the means of fluency metrics show a few interesting findings (Figure 5.6). What shows is that lab session human idle time is on average much longer ($M = 35.17\%$ of total interaction time) than the interactions in formative field sessions ($M = 38.85$ of total interaction time). Field interactions are shorter and likely more task-focused or more frequently interrupted.

	Context Confusion Occurrences	Robot-led Activity Switch Turns	Human idle ratio	ASR Mistakes	Repetitions
Field	0	0	$(M = 0.38, SD = 0.30)$	5 $(M = 1.25, SD = 1.89)$	0

Table 5.2: Fluency metrics (total counts and session averages) from recorded transcripts for field sessions: Context Confusion occurrences counted as the amount of messages, Robot-led Activity Switch Turns, Human idle ratio, ASR mistakes, and LLM repetitions.

5.3.2. Affect Change

Almost all visitors showed a neutral or positive valence emotion during the start of the experiment, and most were positively impacted by the interaction as shown in Figure 5.7. For roughly 33% of the sample, the self-reported emotions suggest visitors were slightly happier (see the violin plots in Figure 5.9 for an indication) after interacting with or seeing the robot. It is interesting to see that most participants' self-reported emotions are to be found in the positive-valence half of the Arousal-Valence space, or increased into that area as can be seen in the right half of Figure 5.7.

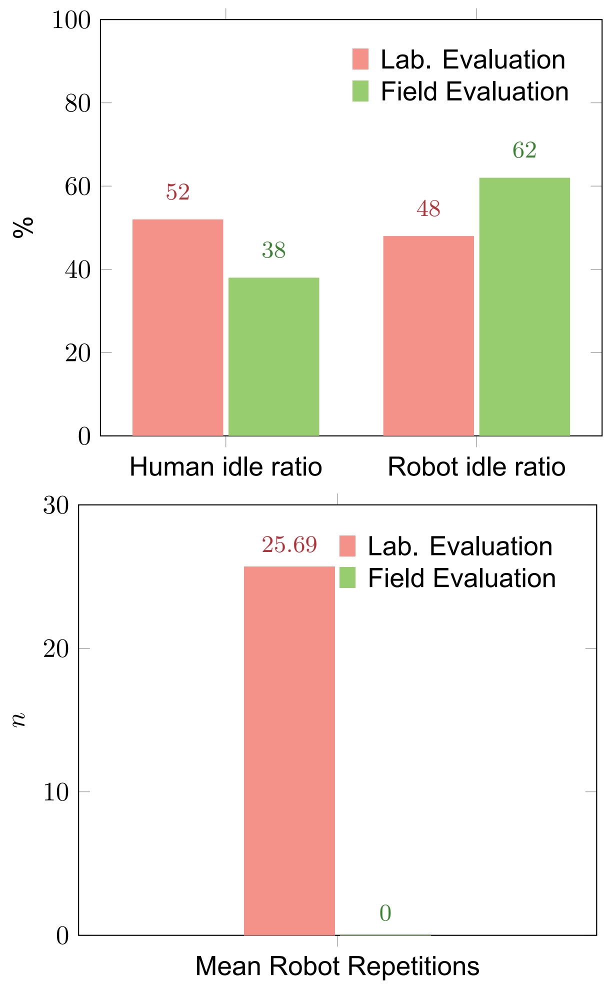


Figure 5.6: Fluency metrics by evaluation setting of the robot interacting with visitors. Top: mean interaction time over total time per interaction (in percentages). Bottom: number of robot utterance repetitions observed per session. Lab = laboratory evaluation; Field = real security constrained waiting-room evaluation.

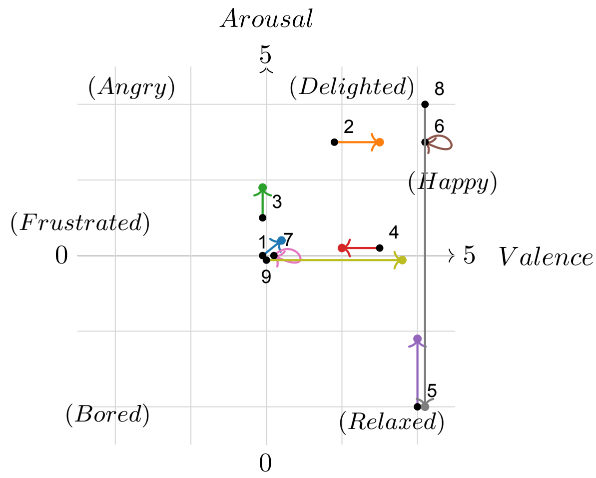


Figure 5.7: Arousal-Valence space of results from the self-assessment Affective slider (Based on a scale of 1-5 where 1 is the most negative and 5 is the most positive). The horizontal axis is Valence and the vertical axis is Arousal. The arrows indicate the difference between pre-interaction and post-interaction emotions of visitors. Note that numbers 6 and 7 did not report any change in arousal-valence scores so their positions do not change from pre to post-evaluation.

5.3.3. Expectation Change

In the waiting room, visitors' pre-interaction expectations were moderately positive ($M = 56\%$ agree) and post-interaction ratings show a notable increase in perceived benefit (Figure 5.8; $M = 89\%$ agreement/neutral combined, with $M = 89\%$ reporting neutral to positive and $M = 22\%$ reporting strong improvement). This suggests that for many visitors, the robot met or exceeded their expectations in subjectively enhancing their visit.

Overall, there was a significant increase in the expectation of visitors towards the robot before the interaction and after the interaction $t(9) = 4, p < .005$ (see Table 5.3, last column).

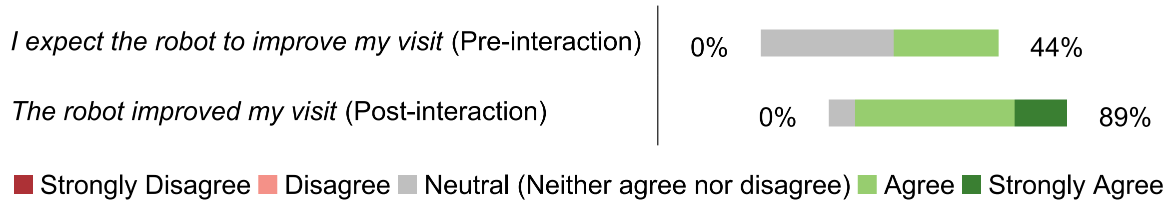


Figure 5.8: Expectation Congruency among field study visitors. The percentages show, respectively, the combined share of responses left of Neutral ((strongly) disagree) and right of Neutral ((strongly) agree). The percentages show, respectively, the combined share of responses left of Neutral ((strongly) disagree) and right of Neutral ((strongly) agree).

5.3.4. Visitor Interactions Field Notes

The following section will present the results from the interactions and interviews in a formative manner. The detailed description can be found in the form of short vignettes below, and a summary is found in Table 5.4. To respect the anonymity of participants the exact order was randomised.

Visitation Perspective 1

The first interaction between the robot and visitors was between a pair of visitors around noon. Both visitors agreed to join the experiment and when the pre-interaction forms were completed only one of the visitors interacted with the robot. Then for four minutes they played a guessing game together. During that guessing game, one visitor discussed potential answers to the game with their spouse. After that interaction, there was enough time before the visit started.

From what they reported afterwards, it holds true that the waiting room time is indeed perceived as boring by most visitors. Apart from the waiting time, it appeared that the clock in the waiting room was only placed recently as they had not seen it before.

Statistic	Valence	Arousal	Expectation
Mean Before	3.92	3.06	3.44
Mean After	4.14	2.78	4.11
Mean change (post-pre)	0.22	-0.28	0.67
SD	0.65	1.43	0.5
t(9)	1.01	-0.58	4
p-value	0.34	0.58	**0.0039**
Cohen's d	-0.36	0.19	-1.33
wsr t(9)	4.0	-2.0	0.0
wsr p-value	0.875	0.31	**0.03125**

Table 5.3: Pre/post study metrics from the field study; means, standard deviations, t-statistics, computed Wilcoxon signed rank (wsr) and corresponding p-values are reported. p-values < 0.05 are shown as significant.

On the note of other matters concerning the physical space of the security constrained waiting room, often there were children brought by other visitors which could be very bored in the waiting area, and they discussed that it is sad that not all visitors might have a sense of association towards each other, which may be attributed to the diversity in backgrounds. That lack of association may mean that other visitors could not be as easy to connect. Apart from other visitors in the room, the atmosphere seemed lacking and discouraging to both visitors for small-talk or entertainment, especially for children or when people want to infer the current time for example.

Both visitors expected that by adding the robot, conversations between visitors themselves (without the robot directly involved) may be initiated that otherwise would not have taken place, just because of the object being in the room. Additionally, they reported that the robot's features and programming (with the guessing game) would better suit to children visiting. One of the visitors said lastly that it was nice that the robot was equipped with a functionality of reporting the time to visitors.

Visitation Perspective 2

At the second visitation hour, another couple entered the waiting room. They reported the waiting was usually quite tedious and long. From their body language and the multiple humorous outcries it appears that they enjoyed the interaction. They played the guessing game with the robot and they took turns telling the robot an answer to the guessing game, until it appeared that the ASR module picked up woman's voices more accurately than man's voices, by the time visitors found out the visitor gave the microphone to their spouse. This interaction was stopped right after they guessed the right animal in the game because the visitation suddenly started. After they returned to the waiting room to exit the building, they answered the other questions and also discussed a bit about the robot.

They both reported that they also remembered the waiting to be rather boring usually. They were enthusiastic about the robot, saying it was something they did not encounter in such a place that often and they reported to want to see the experience more often in the waiting room.

Visitation Perspective 3

Session three took place during the second day. Not all visitors were able to join the experiment due to language barriers. The robot could interact in such a use case with its functionalities but the experiment was not set up properly to adjust for another language.

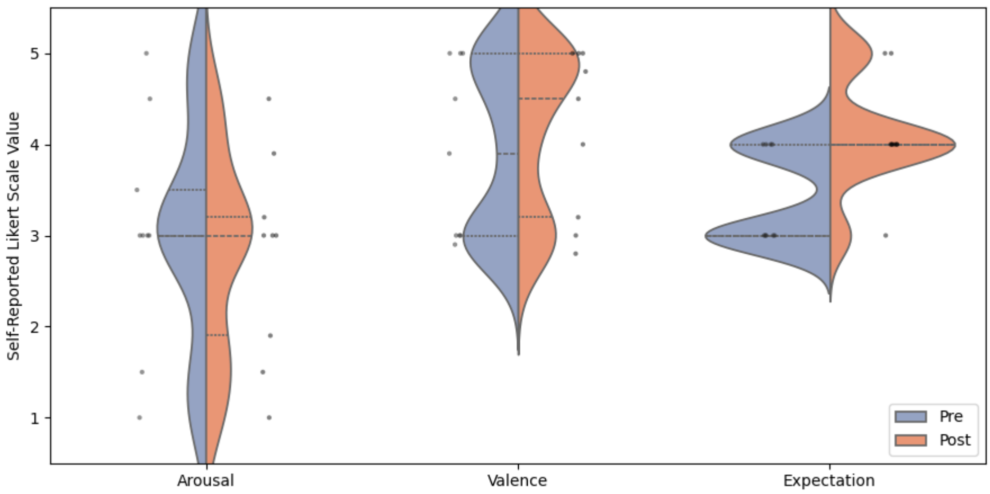


Figure 5.9: Violin plots showing the distribution changes of responses to Arousal, Valence and Expectation before ('Pre') and after ('Post') robot interaction. The left half represents the Pre distribution and the right half the Post distribution; violin widths indicate kernel-estimated density. Inner lines indicate quartiles. The individual points are shown as black jitter points.

Two visitors entered the security constrained waiting room and interacted with the robot while they waited. The interaction made their already giggly state a bit more entertaining. They expressed that they found the robot's appearance endearing saying it was 'very cute!', and they were eager to understand its functionalities.

Another visitor joined two minutes after the interaction was completed between the robot and the other two visitors. This time a visitor brought a younger visitor, who joined the experiment. While the spoken interaction between them only could go on for a minute until the Security Officers called everyone in for visitation, nonetheless the two of them interacted with the robot by letting the little visitor give the robot a physical fist-bump.

It seemed to them that it was necessary that the older visitor helped the younger visitor feel at ease, they reported that it might evoke uncanny reactions from younger people.

Visitation Perspective 4

The last visiting hour was the busiest with visitors. Two visitors joined the experiment during this time slot. They were together with other visitors which made the room atmosphere chaotic. At this time the robot had a technical malfunction and could not be set up in time before the security staff called everyone in for starting the visitation hour.

During this busy waiting moment, it happened three times that participants inquired each other on regulations within the facility. For example they were concerned if they were allowed to introduce self-bought snacks from the vending machine in the corner of the waiting room into the visitation area.

Session	Pp.	Attitude towards robots	Self-reported emotion before vs. after interaction	Pre-interaction expectations	Post-interaction expectations: Did the robot improve your visit?	Interaction Time in min.	Overall
1	V1	neutral	neutral → happier	neutral	positive	4	Animal Guessing Game with turn-taking; reported boring waiting time; reported robot was most suitable for children; suggested robot could spark conversation between visitors
1	V2	neutral	excited → happier	neutral	positive	4	Animal Guessing Game with turn-taking; clock/time reporting appreciated
2	V3	neutral	neutral → a bit more excited	neutral	positive	9	Animal Guessing Game with turn-taking, ASR module picked up this pp.'s voice better, reported boring waiting time, the robot said goodbye
2	V4	neutral	happy	positive	positive	9	Animal Guessing Game with turn-taking, ASR module picked up this pp.'s voice less accurately, the robot said goodbye; enthusiastic about robot presence
3	V5	neutral	calm → more relaxed	neutral	positive	4	Animal Guessing Game with turn-taking, High Five; enthusiastic about robot and found it entertaining
3	V6	neutral	calm	positive	very positive	4	Animal Guessing Game with turn-taking, High Five
3	V7	neutral	neutral	neutral	neutral	1	Fist-bump given to both pp. and other visitor, robot said goodbye; adult pp. expressed concern for uncanny reactions in younger children
4	V8	very positive	delighted → relaxed	positive	positive	3	Busy waiting hour; robot technology was not performing timely; the robot gave a fist-bump to pp.
4	V9	very positive	neutral → happy	positive	very positive	3	Pp. acted very encouraging towards interaction in the same busy hour

Table 5.4: Summary of the results from visitor experiments

5.3.5. Security Staff Perspectives

Semi-structured interviews were conducted at the premises of a real world security facility. 6 employees from distinct roles or departments in the facility were asked to discuss concerns surrounding the potential of a social robot in their work environment.

General current importance or drive for work

Staff members answered the following when asked about what makes their work interesting:

I Employee 1. Stresses importance of personal contact for work and seeing small steps in youth development. Expects that work environment could improve with better work allocation, especially in case of sick leave.

II Employee 2. Also stresses importance of personal contact for work and mentoring youth and seeing their development, albeit in small steps.

III Employee 3. Personal contact is what drives their work.

IV Employee 4. What makes work interesting? Personal contact ('omgang met mensen') but also learning from the diversity of disciplines of staff at the facility. Decision-making about the robot should also focus on integration of the diversity of disciplines ('Meeting with group leaders, guards and management').

V Employee 5. The aim of the work is to add value to society, by improving the lives of the youth.

VI Employee 6. Improving the lives of the youth by letting them find their 'spark': identifying an intrinsic interest, talent, or motivation that can help the youths build up positive development and engagement. Such a spark could support that the youth reconnect to individual meaningful goals, which could improve long-term educational and social outcomes. Seeing and interacting with the robot could close the gap between their current state and the outside world, to ease their rehabilitation.

Helpful Robot functionalities

Staff members answered the following when asked what poses difficulties to what they think is important in their work:

I Employee 1. Stresses that not all functionalities may be implemented due to regulations in the security facility. Wants humans instead of robots taking responsibility in decision-making processes (e.g. opening a door for someone has to be authorized by a human, otherwise it might not be safe). Difficulties of work:

- (a) When youth requires individual attention but not enough employees are available.
- (b) When there is unexpected sick leave and there are work allocations challenges.
- (c) Overuse of the intercom system by youth: the first call overrides others' calls.
- (d) Providing extra pair of eyes (e.g. contraband detection, detection of youth smoking outside which is not allowed for youth below 18) or ears (e.g. if staff walks past a youth that talks about doing drugs inside their rooms, a robot that detects it may be beneficial to staff)

II Employee 2. There is a lot of paperwork which creates work overhead. Difficulties of work:

- (a) Some youth converse in Arabic language, which requires an interpreter. However, interpreters are not always available; remote connections via video-calls are not always stable and create overhead. A locally running robot might provide a better real-time connection and translation.
- (b) Youth struggle with routine and structure, a robotic partner that helps them keep their schedules would be beneficial.
- (c) After an incident youth have to write a reflection paper, but they sometimes struggle with writing. When struggling with writing, the youth can request a conversation but that is also something not all youth feel easy by to disclose their problems. Another functionality for the robot would be augmenting that process by also having a robot that can have a reflective conversation with the youth about their problems and actions, which could provide a way for those who are not good in writing or talking about their issues.

(d) Detection and monitoring: providing an extra pair of eyes (e.g. detecting mood) so that staff can support the youth better.

III Employee 3. Difficulties stem from the fact that staff do not have enough means of detection:

(a) A smell sensor might help staff detect unexpected hazards; for example, a visitor once had drugs hidden in applied nail polish that a youth ingested by licking.

(b) Within the youth living areas vigilance is of utmost importance; sometimes contraband is hidden in the most strange places. Staff expects that a robot detection system may be of support in such cases by providing extra means of monitoring.

IV Employee 4. Robot functionalities:

(a) Monitoring can be done with support of technology, but decision-making should be left at the staff. Recently there was a renovation of the location where technology was added, and the improvement of the amount of camera's has been reportedly helpful to staff.

(b) Becoming a buddy for the youth and providing distraction for the youth with games or music.

(c) Providing information (e.g. instructions during visitation hour). The added value of the robot would be that it can have a disarming appearance, perhaps people who distrust the government have an easier time listening to the robot than to human staff members.

(d) Another functionality would be detection (e.g. searching through rooms for drugs or weapons)

V Employee 5. difficulties at work include:

(a) Detecting contraband on the air yard (e.g. staff may not know when a phone may be hidden and picked up by youth coming on the air yard)

(b) Moving things for staff (e.g. printers are not at the youth living areas).

(c) Administrative tasks (the much dreaded paperwork)

(d) Providing entertainment and information (e.g., on rehabilitation methods or as more secure sandboxed alternative for access to criminal files which is now done on laptops[89]) for the youth

VI Employee 6.

(a) The robot could evoke imaginative thoughts in incarcerated youth at the facility that could benefit their rehabilitation. They explained that with an anecdote about some of the youth wanting to 'become a supermarket cashier' because 'it sounds convenient to them', while they could add more to the societal cause, by possibly finding another calling working in another area such as robotics.

Acceptance of Robots at Work

Staff members answered the following when asked about their opinions on robots in general and of their opinion of the robot deployed at the facility:

I Employee 1. Generally a beneficial attitude towards robots in their work, feels urgency of the current developments only in case of support, not sure if personal contact can be replaced by a robot. Expects the robot to have a too much trusting result on staff so they find themselves cognitively overloading their tasks which may make them less focused or vigilant. Also stresses that the robot should meet expectations (e.g. requiring too much instructions, maintenance or overhead).

II Employee 2. Expects that it will be gradually introducted. Especially likes the Miroki robot which seems friendlier than other robots encountered. Stresses importance of including staff from all layers in the decision making processes of robot platform functionalities: ' because the staff has to work with it not someone in an office'.

III Employee 3. Very positive about robots ('Very open to the possibilities') but acknowledges that humans may be better at detecting intent and lies than robots.

IV Employee 4. Some visitors may be distrustful of institutions and may judge a robot for suspecting the robot to purportedly be harming their privacy.

V Employee 5. The people who maintain the robot system should be part of the early stages of the decision making process.

VI Employee 6. Very beneficial attitude towards robots in their work: The robot could also be a good step for exposing incarcerated youth to the outside world reality, as the youth inside of the facility may be restricted from outside developments.

Effect on Mood and Expectation

Staff members answered the following when asked their expectations of the robot changing the mood of waiting room visitors:

I Employee 1. Games can provide distraction which is especially needed for incarcerated youth who suffer from boredom. That would be better than them causing a fuss when there is no entertainment around. Expects that in case of panic, a human may be better in interventions.

II Employee 2. Thinks its a good idea to add the robot in the visitor area with the animal guessing game. Stresses that the robot should work properly according to expectations, and that it should be of added value.

III Employee 3. Not expecting the mood of the waiting room to change, because the robot will be just another prop to visitors after the novelty effect wears off.

IV Employee 4. The cheerful embodiment of the robot can have a ‘disarming effect’ on agitated people. It reminds of a similar project at the facility: Dutch Cell Dogs [90], where laboratory dogs are trained by the incarcerated youth. It provides distraction of thoughts and improves behavioural problems. If it were to be used with the youth, it may need to be good to make the robot identifiable for them (‘Make the robot joke in street slang’).

V Employee 5. The robot can create a more friendly ambience to the waiting room by providing entertainment. It cannot replace human personal contact.

VI Employee 6 The robot could be an improvement for visitors, especially for lowering anxiety and stress. They explained that officers wear a uniform, which may evoke feelings of anxiety in visitors. Since the robot did not have a uniform, and is more accessible, it would create a different environment for staff, visitors and the incarcerated youth in the facility.

5.4. Discussion

This study investigated how a reception robot can contribute to visitor mood and staff operation within a security facility. The field evaluation tested how real visitors and staff interacted with a friendly robot in the waiting room. Improvements were made on the agent by learning the lessons from the lab experiment.

Compared to the lab, (1) LLM context confusion was mitigated with context engineering, (2) lack of turn-taking and collaboration was improved by addind an animal guessing game, and (3) neutral reception agent functionalities were decomissioned to favour the collaborative activity. This system was evaluated with a different, more friendly robot with real visitors using surveys and with security staff using structured interviews.

The field evaluation aimed to answer the following sub-questions:

I *How can a reception robot contribute to visitor mood and staff operation within a security facility?*

II *Does a collaborative activity (e.g., an Animal Guessing Game) decrease visitor strain and boredom and improve visitors’ overall experience?*

In the real waiting room the robot worked better than in the lab: visitors tended to report more benefit and some became happier after interacting with it. Field visits averaged ≈ 5 minutes with higher turn-taking than lab sessions, so short, active activities suit this setting; the robot should be time-aware, use brief dialogues when transfers are likely, and longer dialogues only when waits are expected (always allowing quick termination).

5.4.1. Findings

Next are the main findings derived from the results:

I **Affect:** Visitors commonly showed neutral-to-positive baseline affect and several participants shifted toward higher valence after interacting with the robot (Figure 5.7); approximately one third reported noticeably happier states. However the responses were variable: while many improved, some showed little change, indicating the intervention often but not uniformly elevated mood.

II **Expectations:** Pre-interaction expectations were moderately positive (Figure 5.8, 56% agree) and post-interaction ratings increased, with 89% reporting neutral-to-positive outcomes and 22% reporting strong improvement. Therefore the prototype met or even slightly exceeded visitor expectations.

III **Fluency:** Field interactions were shorter and more turn-dense than lab sessions (Figure 5.6.a; $M_{field} \approx$ 4.3 min vs. $M_{lab} \approx$ 9.9 min). Human idle time ratio being lower in the field interaction (Figure 5.6.b; $M_{field} \approx 38\%$ vs. $M_{lab} \approx 52\%$) suggests higher conversational density in the adjusted prototype, while robot repetitions decreased completely (See Figure 5.6) after the prototype iteration.

5.4.2. Interpretation

Given the small sample, these results are indicative rather than conclusive. Therefore, the interpretation of the results requires caution but also allows for some inference and indications about visitor reactions and robot design implications.

5.4.3. Limitations

The study has several limitations. The small sample ($n = 9$ visitors, $n = 6$ staff members) and voluntary recruitment limit statistical power and likely bias the sample toward favourable views. This may be exceptional in security-constrained waiting rooms where some visitors may be more sceptical towards institutions and therefore towards taking part in the study. Interactions were short, variable, and often cut short (mean $\approx$ 5 minutes), constraining exposure to the robot. Audio capture and ASR relied on a simple low-cut filter in a noisy environment, risking missed speech (like the visitors not being heard by the module due to less loudness). Staff effects were assessed mainly qualitatively without robust workload or stress metrics, and the measurements are situational and a one-time assessment.

Variation of Visitors' opinions

Not all visitors were available and/or willing to participate in the experiment, which can explain the suggestion that results are mostly favourable towards the robot use in security-constrained waiting rooms. Visitors (particularly if conspiracy thinking is involved) who are showing negative affect during the recruitment of the study in security-constrained waiting rooms may be biased to be reluctant to participate, meaning that a significant amount of opinions could not be used as input for this study.

5.4.4. Recommendations

A mood-enhancing robot deployed in a security-constrained waiting room presents multiple interlinked opportunities for improving visitor experience and operational outcomes.

Recommendations for Co-Design and Testing a Social Robot in High-Risk Domains

We recommend a set of practical steps for co-designing and testing a mood-uplifting robot in a security constrained waiting room. That includes a few recommendations (see Table 5.5 and Table 5.6), varying from making interactions time-aware, prioritizing short, turn-based activities, and building clear failure protocols in collaboration with staff. Also recommended is the continuos iterative involvement of security staff throughout design and testing, using sufficiently large samples to measure effects, and improving technology to ensure the robot is competent.

1. Design for variable time: To make interactions fit the waiting room schedule, one can explore making the robot time-aware. That way, short options (e.g., fist bumps or small talk, greetings) could be available when time is limited, medium-length activities (simple short turn-taking games as the animal guessing game in our research) can be available for average waits, and longer activities (stories or extended play) only when there is ample time. However, the robot should

always allow the interaction to end immediately when visitation begins so relatives don't lose precious time with incarcerated youth.

2. Prioritize conversational fluency and brevity: Focus on smooth turn-taking and avoid long monologues. Small, interactive games that encourage back-and-forth engagement work better than one-sided presentations and help maintain attention and positive mood.
3. Robust failure protocols and staff collaboration: One should develop clear procedures for when the robot freezes or malfunctions, created in collaboration with security staff. These protocols should mitigate risk in all likely scenarios, prevent escalation, and define who intervenes and how, essential in a high-risk environment.
4. Include staff in design and testing: Treat security staff and other on-site workers as developmental partners, not bystanders. Involve them in decision-making, testing, deployment planning, and maintenance planning; take their needs, ideas, and privacy concerns seriously. Design should support their work and preserve their motivation (most of them are quite driven in their work as described in the results section 5.3.5) and operational passion rather than replace or undermine it.
5. Research design and sample size: Use a sufficiently large sample when evaluating the robot so effect sizes can be estimated reliably. Plan for repeated measures or staged rollouts to capture variability across different visitor groups and time conditions.
6. Privacy, respect, and acceptability: Ensure the robot's presence and behaviour respect visitors' dignity and staff routines. Design interactions that are culturally appropriate and sensitive to the emotional context of visitation, and provide easy opt-out choices for visitors who do not want to participate.
7. Technical recommendations: Improve automatic speech recognition to reliably detect different voices in the waiting room and consider multimodal cues (for example, gaze direction) to determine when a person is addressing the robot. Such measures will increase responsiveness and reduce confounding into effects of the robot.

R&D Opportunities and Challenges for a mood-intervention robot

Lastly we outline the Research & Development opportunities and challenges for a mood-intervention robot in the security constrained waiting room. We divide them into the promising opportunities and ideas (entertainment, conversation facilitation, information provision, multi-language support, and a support-buddy for incarcerated youth, see Table 5.5), explain their concrete use cases and benefits, and pair them with the obstacles for development and deployment (see Table 5.6). That way, future teams know what use case to explore, and can anticipate any obstacles to prioritise to make their work impactful.

We first recommend using the robot for the use case that we focused on. We showed that the robot made the waiting room mood uplifting. Visitors and staff liked it, so using the robot this way works well for its purpose in the security constrained waiting area. Visitors reported that the waiting area is boring. A friendly, childlike robot can reduce boredom and lift mood, and that mood can spread to others (mood contagion [13]). Entertainment can be simple: short conversations or games driven by conversational agents backed by LLMs, and light physical gestures such as fist bumps or high-fives. Our presented robot already implements this approach, so applying our research design and recommendations would allow straightforward deployment. We first recommend using the robot for the use case we focused on, since it lifted mood and was liked by visitors and staff, Next, we consider how the robot can also encourage visitors to interact with each other.

Second, our recommendation is that the robot could also encourage visitor-to-visitor conversation. A robot, possibly even an inanimate one, can act as a social catalyst and encourage visitors to talk to each other, which may reduce anxiety in waiting rooms (see [91]). This effect may not require extensive interactivity; only the recommended modifications to the current setup would be sufficient.

Thirdly, while no visitors made use of it, the staff reported that they found that visitors often had informational needs, highlighting the purpose of the robot as information provider in the visitor area. During busy waiting periods, participants repeatedly asked each other about facility rules (for example, whether snacks bought from the vending machine are allowed in the visitation area) and about simple facts like the time. This shows a clear informational need. A robot could answer routine questions, collect statistics about common information requests for staff, and assist supervision by flagging contraband-related questions.

Fourth, what is not explored extensively in the field evaluation is multi-language support, which nonetheless may be impactful in future applications of social robotics in the high-risk domain. Many visitors speak languages other than the facility's primary language, as discussed in visitor vignettes and staff interviews. Adding Arabic, Turkish, or other language support would increase the robot's usefulness in various situations. Technically this requires modest parameter changes and user testing with native speakers to ensure quality, since some languages are easier for LLMs to handle than others. The original experiment did not include this adjustment, but it would add substantial value with limited engineering effort.

Lastly, an interesting area for social robots in the high-risk domain will be that of a support buddy for incarcerated people. Based on perspectives from staff, the robot could serve as a private conversation partner to help people write reflection papers or discuss incidents when they feel uncomfortable talking to staff or peers. This could help those who struggle with writing or disclosure and give staff insights into whether youth are learning from mistakes. Challenges to that deployment include peer stigma (the 'inmate code') that may discourage group use, but it does not stand in the way of personal 1-to-1 use of the buddy robot. Then there are challenges to privacy and safety requirements, especially within the legal possibilities of the security constrained domain, and the need for extensive adaptation and user testing to prevent malfunctions and protect sensitive data.

Title	Possible use cases	Motivation / Pros
Disarming entertainment by waiting-area engagement	Waiting room area; small conversations; simple games; physical interactions (fist bumps, high-fives)	Visitors report boredom; improved mood via robot can spread to others (mood contagion [13]) and reduce perceived waiting unpleasantness; leverages existing prototype and recommended design for easy implementation
Encouraging visitor-to-visitor conversation	Passive/inanimate robot presence or minimally interactive robot to act as social catalyst	Can increase social interaction and reduce anxiety in waiting rooms (effects of inanimate robots are reported in prior studies e.g. [51]); may require little change from current setup
Information provision in visitor area	Answer common questions (rules, visiting regulations, time); collect aggregate data on informational needs; assist supervision (flagging contraband-related questions)	Fulfils clear informational need observed in field; reduces staff burden; provides data for staff training and oversight
Multi-language support	Interact with visitors in Arabic, Turkish, or other languages; provide translated information and simple conversation. Useful in many use cases across the facility, e.g.: translation between visitors, incarcerated youth and staff, picking up discussions in languages that staff is not equipped to listen to, interactions in visitor area.	High added value for non-native speakers; modest technical changes (parameter/content updates) and targeted user testing with native speakers can validate performance
Support buddy for incarcerated youth (reflective conversation aid)	Help youth with reflection on incidents and actions; provide private reflective conversations after incidents; scaffold writing and self-reflection	Supports rehabilitation tasks, may be a support for those uncomfortable disclosing to humans, and can augment staff insight into learning progress; aligns with the facility's societal objectives

Table 5.5: Opportunities for a robot in the security constrained facility: idea, use cases, and motivation/pros. Based on visitor and staff perspectives.

Challenge	R&D community challenges	Motivation / Why it's challenging
Safe, engaging entertainment design	Designing LLM-driven dialogue that is brief, appropriate, and resilient to adversarial/user misuse; integrating reliable physical gestures (fist bumps)	Need to balance safety, robustness, and social appeal; physical interactions raise safety and reliability requirements; content moderation and context-awareness required
Eliciting authentic visitor interactions	Measuring and reliably producing social catalysis with either active or passive robots; determining minimal features needed	Social effects are context-dependent; effects may vary by population and environment, so reproducible experimental designs and longitudinal studies are needed
Accurate, audience-appropriate information provision	Ensuring clear answers for target audience; handling ambiguous or legal questions	Risk of giving incorrect or legally sensitive guidance; requires maintenance pipelines with staff in the loop, up-to-date content sources (such as [79]), and careful logging/consent design
Multi-language Support Robustness	Achieving natural, culturally appropriate interactions across languages; validating translations and dialogue nuance	Some languages/translations degrade LLM performance; requires native-speaker testing, localised prompts, and possibly specialized models or fine-tuning
Privacy, trust, and group-dynamics for youth support	Protecting privacy, avoiding 'punitive' surveillance, overcoming peer stigma ('inmate code'), and ensuring clinical appropriateness	High-risk setting: needs strict privacy-preserving design, safe practices and protocols for inmates' sensitive disclosures, and strategies to address social acceptability among youth

Table 5.6: Challenges: technical and social barriers and why they matter. Based on visitor and staff perspectives in the security facility. The technical and social challenges are ranked roughly by increasing difficulty.

6

Discussion & Conclusion

This study examined the impact of a reception robot on visitor mood and staff workflows at a Dutch juvenile detention facility. While earlier studies have explored the impact of social robots on mood intervention in health care (e.g., [70, 1, 54]), they have not explicitly addressed its potential effect in the security constrained domain. An iterative approach [26] was used to prototype an initial reception robot system for the security constrained waiting room. The design was divided into an initial prototype phase (see Design Rationale in chapter 3), a lab evaluation phase (chapter 4) and with improvements from the previous evaluation an improved system was made which was evaluated in a field evaluation phase (see chapter 5).

6.1. Key Findings

Where the first iterative prototype was criticized for its interaction style (e.g., long monologues, poor turn-taking), the second iteration was received positively, suggesting it met expectations and improved visitation waiting experiences whilst allegedly gaining acceptance from security staff. The results indicate that the provided adjustments (i.e., context prompt management, animal guessing game, removal of functions that increased monologues) helped build a prototype design which adequately suits itself for interaction within its domain.

I *How does a informative robot receptionist welcome visitors, inform them about rules and appointment times?*

In the laboratory evaluation, the storyteller prototype produced larger affective shifts than the informative receptionist (storyteller > receptionist; see Figure 4.6), but those shifts were mixed: some participants improved markedly while others moved toward negative valence. Expectations declined more after storyteller interactions than after information provision interaction, (see Figure 4.9). In the field evaluation, informative functionalities were not used by visitors even though most visitors revealed an unmet informational need (i.e., visitors asking each other; *'Can I take a bought vending machine snack into the visitation room?'*, see section 5.3.4).

II *Does a backstory fantasy robot persona decrease visitor strain and boredom and improve the visitors' overall experience?*

It may be unclear what the storyteller robot can do for visitors within their waiting time constraints in the real security facility. However, the lab evaluation experiment indicates that the backstory fantasy robot persona did not receive the expected result. The laboratory prototype limitations (low conversational fluency, see Figure 4.5) and experimental limitations (small sample, ecologically less valid lab setting) constrain generalisability and likely amplified negative responses for a subset of participants.

III *Does a collaborative activity (e.g., Animal Guessing Game) decrease visitor strain and boredom and improve visitors' overall experience?*

The security constrained waiting room field evaluation robot with its iterative adjustments produced clearer positive signals on mood and overall expectation. Visitors entered the interaction with favourable expectations that the robot would improve their wait; alike the lab experiment participants (both Figure 4.8 and Figure 5.8 show more than 89% visitors self-report neutral or positive expectation pre-interaction); post measures in field expectation change were positive ($t(9) = 4, p = .0039$, Figure 5.8), indicating the improved prototype with animal guessing game met or modestly exceeded anticipated benefits for most visitors. Furthermore, most field evaluation visitors reported neutral-to-positive valence at baseline and shifts toward more positive valence after interacting with the robot; however not exhaustively significant, one third does suggest a slightly happier (suggested increase of Valence from $M_{before} = 3.92$ to $M_{after} = 4.14$, $t(9) = 1.01, p = .34$) state after the robot compared to before interacting with it. This may suggest that short, low-effort collaborative interactions may produce measurable mood changes in a security constrained waiting-room setting. The more positive overall experience can be explained with the fluency metrics: field interactions were shorter and had less human idle time compared to the lab evaluation (see Figure 5.6), visitors more often shifted toward higher valence (about one third noticeably happier, see Figure 5.7) and post-interaction expectations improved for mostly all participants (see Figure 5.8). While no direct causality can be derived from the results, the lab experiment differs vastly in sample and setting from its ecologically more valid follow-up in the real security constrained waiting room, fluency indicators (human idle time and repetitions) appear to be lower in the second iteration of the prototype (compared to that of the laboratory examination), suggesting an improvement in the turn-taking. That improvement is reflected with a positive acceptance reported by visitors: smoother interactions with less human idle time coincided with more positive affect for visitors.

IV *How can a social reception robot with supervision capabilities support staff operation?*

Staff preferences for simple, easy-to-learn, controllable functions indicate the importance of practical, accessibly designed features (e.g., staff-triggered messages, clear override controls) for the adaptation by staff. Functionalities were also seen as potentially helpful when explicitly delegated and easily overridden by staff, such as operator supervision and control override for the context modes that were incorporated in the design. The data implies operational value when robot actions are transparent and subordinate to the needs and wants of staff members. Therefore, the security staff may find a benefit from the knowledge building that is done as a by-product of this research. Staff report that they did value personal contact as one of their important reasons of their work, but they were also enthusiastic about the robot supporting in the waiting room and possibly improving their workplace atmosphere.

Differences between lab and field outcomes are plausibly explained by ecological realism and interaction length: lab sessions were longer and more monologic, yielding higher idle time and unrealistic expectations, while field sessions were brief and naturally constrained, favouring short, active exchanges. Iterative improvements (ASR, dialogue edits, time-aware behaviour) reduced repetitions and improved turn-taking, which likely accounts for better field fluency. Where experiments disagree (e.g., magnitude of affect change), treat lab signals as mechanistic hypotheses and field results as preliminary evidence of external validity.

The lab and field studies both considered, this thesis found indications supporting a personality-driven, mood intervention robot that can influence visitor affect and expectations in the security constrained waiting room, but outcomes depend strongly on the design of the mood intervention capability and the robot fluency.

6.2. Interpretation of Results

Our study suggests that the security constrained waiting room visitation space is a promising research setting for robots as mood- or stress-intervention tool, while opening a vast realm of operational functionalities. This aligns with a growing body of research deploying reception robots in care sectors, and deploying technology for mood intervention such as to reduce vaccination anxiety[1], blood draw[54], support dental procedures [14], and provide distraction for people with dementia [55].

Although de Jong et al. [6] suggests problem-solving topics are disadvantageous to overall visitor affect, our findings suggest a more nuanced effect when robots intervene: visitors interacting with a social robot may actually benefit from a joint problem-solving activity like an animal guessing game. What might explain that is de Jong et al. [6]'s method which focuses on conversations *within* the visitation hour, whereas our method focuses on the conversations with the robot *prior* to the visitation hour. While we do not collect data from during precious interaction of visitors and the incarcerated individuals they visit, we did gather data from visitors after their interaction.

Staff feedback identifies specific improvements that increase the robot's acceptance and value within future steps of the iterative process. Most notably, staff members subjectively reported that direct interpersonal contact is the most valuable part of their work, but they viewed the robot as a useful help because it could provide an additional means of observation. That means of observation could be implemented as an added functionality which can be found in the works of Siegmann and Bendel [10].

Lastly, the prototype supports a privacy-centred approach that is wanted in a security constrained waiting room located in a country upholding its democratic values that it supports. The prototype adopts a privacy-centred approach that aligns with a democratic and humane security facility. This design prioritizes data collection minimization and local processing to reduce the collection and transmission of sensitive information. Staff and administrators expressed preference for these measures, citing legal and ethical requirements as well as the need to preserve inmates' dignity. Eventually, we expect the prototype's privacy features to increase institutional trust and facilitate acceptance, making it more likely to be adopted and integrated into routine practice.

6.3. Addressing Limitations

This study explored the iterative deployment of a social robot in the waiting room of the security constrained waiting room focusing on mood-intervention activities. The findings suggest the deployment received positive acceptance by visitors and staff members, confirming the uplifting role of a mood-intervention agent. However, long-term studies may be needed to confirm deployment effects, especially regarding decision-making, embodiment and technological maintenance.

6.3.1. Potentially Masking Systemic Problems

One pitfall of deploying a social robot in a security constrained waiting room is that it risks being a technological fix to social problems[53]. As such, the robot we designed may amplify feelings of surveillance, loss of control, and depersonalization by addressing surface-level interactions without changing the underlying power structures, procedures, or emotional burdens that drive distress. One should make sure that particular innovations are necessary, and always consider alternatives, therefore we want to provide alternatives in the next paragraphs.

Similar to some environmental changes like media playing in a waiting room [51], a robot does not change institutional procedures, and may even reinforce feelings of surveillance or depersonalization. The robot could then end up functioning as a coping layer for deeper systemic issues inherent to the institutions of courts, penitentiaries, airport or immigration customs. Most of these spaces will inherently provoke anxiety because their necessary functions are assessing threats, restricting freedom, surveillance, power asymmetry, and making high-stakes decisions. Those may intrinsically produce perceived loss of control, uncertainty, emotional burden, and social-evaluative threat.

While the robot should be seen as part of measures reducing avoidable stressors, researchers and policy-makers should deliberately pursue alternative environmental, organizational, human-centred interventions. Structural factors may require domain-specific structural changes. For example, systemic improvements in the perceived fairness of domains like immigration, penitentiaries and law can signal fairness to the broader community (visitors and family), whose moods are psychologically tied to the perceived justice of the system[92]. Rehabilitation and improving the chances of incarcerated people in society will increase theirs and their visitors' mood, because a main objective of the support network may be the successful social reintegration of the incarcerated individual, structural investments in rehabilitation serve as 'hope-generating mechanisms' promoting mood[93]. A human-centric approach would foster better mood in the lives of visitors to incarcerated people by providing clear pathways to a productive post-release life, shifting the focus of the visitation from a site of shared hardship into a site of vicarious agency. Future designs must consider whether technology acts as a band-aid or a genuine tool for navigating these burdens.

6.3.2. Sampling and Generalisability

Both experiments used small samples that limit statistical power and the external validity of the findings. The laboratory study $(n = 17)$ relied on convenience recruitment by the researcher; although the lab sample included some demographic variety, it likely over-represents participants who are comfortable engaging in university settings. The field study used a small, voluntary sample of visitors ($n = 9$), which creates additional self-selection bias: visitors who agreed to interact with the robot may be more open to technology or more positively disposed toward staff and institutional actors than the average visitor. This phenomenon may be more pronounced in the circle of security constrained facility visitors, based on information found in interviews with staff members. Together these sampling constraints mean effect sizes reported here should be treated as preliminary estimates; larger, more diverse samples are needed to assess heterogeneity across age, education, language background, and levels of visit-related stress.

6.3.3. Lack of Adaptive Interaction Activities

The two studies focused on distinct designs for the interaction roles. Instead of preferring either a long form content robot or a short form content robot, the robot could be made time-aware and supply content based on the available time. Because sessions in the lab were of different duration than field interactions, a single dialogue system that dynamically selects content length and interaction role based on remaining wait time would perform differently. This limitation matters for both effectiveness and safety: as discussed in chapter 4, with the narrative or information provision role, the robot risks delivering long monologues that frustrate users with limited time (increasing negative affect and idle time), or conversely offering only trivial short games when deeper engagement would be appropriate. It may also constrain operational deployment in the security constrained waiting room: a non-adaptive robot may interrupt visitation hours or fail to react appropriately when supervision demands it, creating more staff burden overhead. To address this, future iterations of social robots in the security constrained waiting room should implement and evaluate a time-aware policy (e.g., short greetings and mini games for > 5 minutes, medium activities (e.g. information answering) for 5 to 15 minutes, extended stories for > 15 min) and measure outcomes across varied wait distributions, including metrics for fluency, user affect, and staff acceptance.

6.3.4. Increasing Societal LLM Acceptance

The recent developments and popularity of LLM chatbots across society may may have shaped participants' expectations and novelty responses. However, the robot's embodied nature; physical presence, arms/legs, facial expressions, and proxemics, fundamentally differs from disembodied conversational agents and we assume it produces its own original social and affective responses. These two sources of influence (general LLM enthusiasm vs. unique embodied cues) are confounded in our data: participants' favourable comments could reflect novelty or LLM 'hype', while embodied cues may be responsible for mood contagion and non-verbal engagement. Future research may investigate whether the effects measured within the security constrained domain by our study will require controlled comparisons between embodied versus screen-based agents.

6.3.5. Social Desirability

We acknowledge that while our methodology aimed to capture valid data both in lab and field study, participants still may have exhibited response bias. Perhaps in some cases of our research, a robot would "steal the show", even though the research focus was to learn about everyday life with a robot in the security constrained waiting room. To avoid appearing rude or offending the researchers, visitors may have inflated favourable feedback and obscuring their true attitudes and/or mood towards the robot and the interaction. The effect of this is unknown to us, but could become known by randomizing which sessions have the experimenter visible versus hidden, and analysing their effects. Future protocols may build upon our methodology of gathering anonymous responses and indirect questioning and may expand by or counterbalancing experimenter presence.

	Requirements	Rationale
1	The robot should respect visitors' time by adapting its activity (information answering, storytelling, gaming) to the visitor.	Field study results show better experience ($t(9) = 4, p = .004$) than lab study experience ($t(17) = -1.19, p = .251$).
2	The robot should be time-aware: estimate remaining available interaction time also useful for proactively tailoring content length and complexity.	Lab and field study results show that participants and visitors appreciate a trivial time telling functionality in spaces where their mobile devices and watches are prohibited.
3	The robot should use short speech acts (micro-utterances) constrained to 5 seconds/a long sentence and combine them into longer content only when user engagement permits.	Lab study results suggest low experience and valence with the storyteller robot, anecdotally due to lengthy robot monologues, high human idle time and robot repetitions, see Table 4.4.
4	The security waiting room robot should take note of using locally run LLMs for data locality.	Design and staff interviews, see section 5.3.5.
5	Robust prompt-engineering templates to constrain behaviour and reduce context confusion.	Lab study results indicate repeated LLM context confusion in robot output creates unwanted interactions, field study results eliminated such behaviour and was positively accepted: see Table 5.2.
6	The robot must provide language support for visitors not fluent in the site's primary languages.	Field study and staff interviews, see subsection 5.3.5

Table 6.1: Requirements for a social robot in security waiting rooms and their rationale.

6.3.6. Evaluation of Staff Training and Maintenance

Apart from an exploratory, interview-based approach, not all practical operational constraints were tested exhaustively due to the short amount of time and the scope of the study. Practical constraints such as privacy, safety, staff roles in co-working and maintenance of the robot, remain open barriers to deployment. These are promising for future work, because it remains important to make sure technologies are sustainably maintainable. In future studies on the research of social robots in the security constrained waiting room, one can learn from our findings for staff members.

6.4. Implications for Future Work

Narrative, personality-rich interactions have higher potential impact than neutral information delivery but are riskier: without careful turn-taking and context management they can produce disappointment or frustration for some users. Short, actively engaging activities that encourage frequent turns are better suited to real waiting-room conditions than long monologues. Technical improvements are necessary to improve conversational fluency; activity-specific context prompts, stronger memory pruning, reduced on-device model constraints, and improved ASR filtering. Field deployments must prioritise simple staff controls and transparency to support uptake and operational fit.

Both evaluation prototypes used a local model set up. Local (on-device) models are promising for human–robot interaction in high-risk, privacy-sensitive settings (e.g., penitentiaries), but hybrid approaches suggest implications for a higher fluency. In our study we did not explicitly observe quality drops of LLM-generated content in our study; it is unclear yet since we did not test (i.e. with an A/B experiment) a cloud-API LLM against locally run models, so the impact on robot fluency remains uncertain. Zhang et al. [18] agrees to that when they describe how local models can produce lower-quality or less fluent outputs than cloud models; users often trade privacy for generation quality and speed, but hybrid approaches can balance privacy and performance. Another obstruction to using local models is the economics: the RTX 9050 used for our evaluation is currently priced at $4000 or €3300. In the future, research efforts need to address these challenges to give industries affordable privacy (data and/or model locality) and performance. For example, PrivateLoRA preserves data locality while cutting communication overhead by >95% and achieves high throughput and tuning performance for deployment [94]. We recommend evaluating such a hybrid setup (sensitive processing on local models, heavier encrypted generation via cloud or offload) and run A/B tests against a cloud API to quantify fluency.

The lab experiment (not conducted in the actual visitation room) yielded less positive results than observations from the real waiting area: participants sometimes found it difficult to have informative conversations and perceived interactions as too long. In the real visitation setting the experiment was better received; joint activities and more open conversation appear useful to increase engagement. Prior work also indicates the need for improved turn-taking functionalities for a fluent conversational robot.

The ASR module used a simple low-cut filter; speakers with soft voices may therefore be poorly recognized. The literature recommends beam-forming as a more robust solution (see, e.g., Haeb-Umbach et al. [95]).

Increased monitoring necessarily elevates the invasiveness of surveillance, which can adversely affect visitors' affective states and expose facilities to reputational and ethical controversy if not carefully managed. For example, the Hawthorne effect predicts that individuals who perceive themselves to be observed will alter their behaviour [96], which may confound measured responses to the robotic system. In recognition of privacy concerns, our study used transparent procedures. We explicitly described the types of data collected and obtained informed consent prior to each interaction. To reduce inadvertent capture of non-consenting visitors, we implemented a hardware filter on the robot's microphone that limited audio input to a radius of approximately one metre. These measures mitigated but did not eliminate the ethical and practical challenges associated with increased surveillance. Therefore, important questions remain for future deployments. How should privacy, consent, and acceptable levels of monitoring be regulated in public-facing robotic systems? Addressing these issues will require attention from legislators and policy makers to develop clear standards that balance visitor well-being, staff safety and public trust.

For future work, we recommend a staged path:

1. Implement the fluency and ASR fixes and redesign activities to favour rapid exchanges and add multilingual support for the contextual domain
2. Implement the conversational agent with a full-duplex model like NVIDIA PersonaPlex [97] so that conversations have natural turn-taking, handles interruptions, backchannels and natural conversation rhythm.
3. Conduct a larger, more diverse, longitudinal field study in multiple security constrained waiting room locations to assess long-term effects on visitor mood and staff workplace atmosphere improvement
4. Collect richer, continuous measures (longer session sampling, objective behavioural markers, and stress/workload metrics for staff such as the NASA-TLX [98]) to establish statistically significant causal links between mood intervention robot design and affective outcomes.

With these steps, a storytelling robot can be refined from an intriguing prototype into a reliably beneficial waiting-room companion.

6.5. Conclusion

We presented a social robot design for security-constrained waiting rooms that addresses visitors' anxiety and low mood while supporting staff situation awareness. We examined (i) which robot roles and LLM-based dialogues foster a more pleasant waiting environment and (ii) what information can support supervisory staff.

For the first part of this research question, the design operationalized three complementary roles (receptionist, storyteller, and playmate) within an integrated LLM-based interaction framework. First, the receptionist and storyteller roles were implemented in an initial prototype and evaluated in a laboratory study. Although the concept proved feasible, the evaluation identified limitations in the robot's speech acts, including slow responsiveness, overly lengthy utterances, and insufficient timing. These findings informed a design iteration in which the interaction was refined and extended with a playmate role and monitoring function. The resulting second prototype was evaluated in a field study conducted in a secure waiting room of a public service facility.

Our findings confirm that non-threatening social robots can significantly improve visitor experiences in security constrained waiting rooms. Specifically, interactions with engaging applications, like an animal guessing game, were favoured over reception and storytelling agents in controlled environments. This indicates that the robot's context and identity are crucial for user acceptance, as real-world interactions often provide more relatable engagement.

Interviews with staff revealed that new potential roles for social robots in security settings can be identified from the staff interviews: the reception playmate could be a successful application, added supervision was reportedly desirable, and additionally, the robot could function as an interpreter or buddy, to further assist in facility operation. This versatility highlights their potential to support both visitors and staff. By combining laboratory and field tests, we ensure that insights gained in controlled settings translate effectively to real-world applications, ultimately promoting a more supportive and humane facility environment.

It is expected that in the near future we may see friendly mood-intervention robots deployed in security constrained waiting area, or even beyond that in other places within security facilities. That may be either as mood-lifting companions and information providers, but also as multilingual interpreters or private support buddies for incarcerated youth.

Bibliography

[1] A. Neerincx, T. Hiwat, and M. de Graaf, "Social robot for health check and entertainment in waiting room: Child's engagement and parent's involvement," in *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, ser. UMAP '21. ACM, Jun. 2021, p. 120–125. [Online]. Available: http://dx.doi.org/10.1145/3450614.3463413

[2] H. Lee, N. Cila, K. Koerten, M. Bleeker, and M. Rozendaal, "The human-robot ensemble: Designing expressive movements for restaurant service robots," in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '25. IEEE Press, 2025, p. 312–320.

[3] J. M. Rodríguez-Pellejero, I. Mulero-Henríquez, and Z. Santana Amador, "Real-time stress monitoring in a child-friendly court: a repeated measures field study," *Humanities and Social Sciences Communications*, vol. 11, no. 1, pp. 1–10, 2024.

[4] K. Konya and A. Sieben, *Royal Society open science*, vol. 10, no. 6, p. 221601, 2023.

[5] R. Triebel, K. Arras, R. Alami, L. Beyer, S. Breuers, R. Chatila, M. Chetouani, D. Cremers, V. Evers, M. Fiore, H. Hung, O. A. I. Ramirez, M. Joosse, H. Khambhaita, T. Kucner, B. Leibe, A. J. Lilienthal, T. Linder, M. Lohse, M. Magnusson, B. Okal, L. Palmieri, U. Rafi, M. van Rooij, and L. Zhang, *SPENCER: A Socially Aware Service Robot for Passenger Guidance and Help in Busy Airports*. Springer International Publishing, 2016, p. 607–622. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-27702-8_40

[6] E. W. A. de Jong, H. Palmen, A. A. T. Ramakers, and P. Nieuwbeerta, "Unraveling the black box of prison visitation: Incarcerated individuals' and visitors' conversations and feelings during visitation hour," *Crime & Delinquency*, vol. 70, no. 6–7, p. 1783–1808, Sep. 2022. [Online]. Available: http://dx.doi.org/10.1177/00111287221125388

[7] P. Serail, "Volle bak," accessed on 06-05-2025. [Online]. Available: https://www.nemokennislink.nl/publicaties/volle-bak/

[8] F. Weerwind, "Brief van de minister voor rechtsbescherming," Tweede Kamer der Staten-Generaal, 24 587 Justitiële Inrichtingen, Nr. 926, Den Haag, aan de Voorzitter van de Tweede Kamer der Staten-Generaal. In mijn brief van 26 oktober 2023 heb ik uw Kamer geïnformeerd.

[9] H. Mahdi, S. A. Akgun, S. Saleh, and K. Dautenhahn, "A survey on the design and evolution of social robots — past, present and future," *Robotics and Autonomous Systems*, vol. 156, p. 104193, Oct. 2022. [Online]. Available: http://dx.doi.org/10.1016/j.robot.2022.104193

[10] T. Siegmann and O. Bendel, *Social and Collaborative Robots in Prison*. Springer Nature Singapore, 2025, p. 432–445. [Online]. Available: http://dx.doi.org/10.1007/978-981-96-3522-1_37

[11] B. Crewe, J. Warr, P. Bennett, and A. Smith, "The emotional geography of prison life," *Theoretical Criminology*, vol. 18, no. 1, p. 56–74, Sep. 2013. [Online]. Available: http://dx.doi.org/10.1177/1362480613497778

[12] O. Bachar and J. Guetzkow, "Prison visits and inmates' emotions: a pretest-posttest study," *Journal of Experimental Criminology*, vol. 19, no. 3, p. 809–833, May 2022. [Online]. Available: http://dx.doi.org/10.1007/s11292-022-09514-0

[13] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, "Mood contagion of robot body language in human robot interaction," *Autonomous Agents and Multi-Agent Systems*, vol. 29, no. 6, p. 1216–1248, Aug. 2015. [Online]. Available: http://dx.doi.org/10.1007/s10458-015-9307-3

[14] Y. Kasımoğlu, S. Kocaaydın, □. Batu, G. İnce, and E. B. Tuna-İnce, “The impact of a humanoid robot on children’s dental anxiety, behavior and salivary amylase levels: A randomized clinical trial,” *The Journal of Pediatric Research*, vol. 10, no. 2, p. 132–141, Jun. 2023. [Online]. Available: http://dx.doi.org/10.4274/jpr.galenos.2023.01979

[15] M. Kling, A. Haeussl, N. Dalkner, F. T. Fellendorf, M. Lenger, A. Finner, J. Ilic, I. S. Smolak, L. Stojec, I. Zwigl, and E. Z. Reininghaus, “Social robots in adult psychiatry: a summary of utilisation and impact,” *Frontiers in Psychiatry*, vol. 16, Feb. 2025. [Online]. Available: http://dx.doi.org/10.3389/fpsyt.2025.1506776

[16] A. Łukasik and A. Gut, “From robots to chatbots: unveiling the dynamics of human-ai interaction,” *Frontiers in Psychology*, vol. 16, Apr. 2025. [Online]. Available: http://dx.doi.org/10.3389/fpsyg.2025.1569277

[17] M. Allouch, A. Azaria, and R. Azoulay, “Conversational agents: Goals, technologies, vision and challenges,” *Sensors*, vol. 21, no. 24, p. 8448, Dec. 2021. [Online]. Available: http://dx.doi.org/10.3390/s21248448

[18] C. Zhang, J. Chen, J. Li, Y. Peng, and Z. Mao, “Large language models for human–robot interaction: A review,” *Biomimetic Intelligence and Robotics*, vol. 3, no. 4, p. 100131, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667379723000451

[19] S. T. Fiske, “Stereotype content: Warmth and competence endure,” *Current Directions in Psychological Science*, vol. 27, no. 2, p. 67–73, Feb. 2018. [Online]. Available: http://dx.doi.org/10.1177/0963721417738825

[20] M. M. Scheunemann, R. H. Cuijpers, and C. Salge, “Warmth and competence to predict human preference of robot behavior in physical human-robot interaction,” in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Aug. 2020, p. 1340–1347. [Online]. Available: http://dx.doi.org/10.1109/RO-MAN47096.2020.9223478

[21] J. Moses and G. Ford, “See spot save lives: fear, humanitarianism, and war in the development of robot quadrupeds,” *Digital War*, vol. 2, no. 1–3, p. 64–76, Nov. 2021. [Online]. Available: http://dx.doi.org/10.1057/s42984-021-00037-y

[22] J. POSNER, J. A. RUSSELL, and B. S. PETERSON, “The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology,” *Development and Psychopathology*, vol. 17, no. 03, Sep. 2005. [Online]. Available: http://dx.doi.org/10.1017/S0954579405050340

[23] Boston Dynamics, Inc., *Spot®: Instructions for Use. Safety and Operations Manual*, Copyright © 2024 Boston Dynamics, Inc., 200 Smith St., Waltham, MA 02451, USA, jun 2024, original Instructions. [Online]. Available: https://support.bostondynamics.com

[24] Sony Group Corporation, *Help Guide: ENTERTAINMENT ROBOT ERS-1000*, Copyright 2018 Sony Group Corporation, 2018, 4-724-607-21(6). [Online]. Available: https://www.sony.com/electronics/support

[25] SoftBank Robotics, *PEPPER Handbook*, Copyright © 2017 SoftBank Robotics, 2017, support: 1.844.737.7371; https://www.softbankrobotics.com/us/contact.

[26] E. Saad, J. Broekens, and M. A. Neerincx, “An iterative interaction-design method for multi-modal robot communication,” in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Aug. 2020. [Online]. Available: http://dx.doi.org/10.1109/RO-MAN47096.2020.9223529

[27] ——, “A little chit-chat goes a long way: Design and evaluation of task-and person-oriented styles for social robots,” in *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*. IEEE, Aug. 2024, p. 1712–1719. [Online]. Available: http://dx.doi.org/10.1109/RO-MAN60168.2024.10731475

[28] A. Tuomi, I. P. Tussyadiah, and P. Hanna, “Spicing up hospitality service encounters: the case of pepper™,” *International Journal of Contemporary Hospitality Management*, vol. 33, no. 11, p. 3906–3925, Feb. 2021. [Online]. Available: http://dx.doi.org/10.1108/IJCHM-07-2020-0739

[29] R. Wolter, K. V. Hindriks, D. Samur, and C. M. Jonker, *A Study on Automated Receptionists in a Real-World Scenario*. Springer International Publishing, 2020, p. 340–352. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-49778-1_27

[30] Z. A. barakeh, S. alkork, A. S. Karar, S. Said, and T. Beyrouthy, “Pepper humanoid robot as a service robot: a customer approach,” in *2019 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART)*. IEEE, Apr. 2019. [Online]. Available: http://dx.doi.org/10.1109/BIOSMART.2019.8734250

[31] J. Niu, C.-F. Wu, X. Dou, and K.-C. Lin, “Designing gestures of robots in specific fields for different perceived personality traits,” *Frontiers in Psychology*, vol. 13, Jun. 2022. [Online]. Available: http://dx.doi.org/10.3389/fpsyg.2022.876972

[32] H. Brabra, M. Baez, B. Benatallah, W. Gaaloul, S. Bouguelia, and S. Zamanirad, “Dialogue management in conversational systems: A review of approaches, challenges, and opportunities,” *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 3, p. 783–798, Sep. 2022. [Online]. Available: http://dx.doi.org/10.1109/TCDS.2021.3086565

[33] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, “Retrieval-augmented generation for knowledge-intensive nlp tasks,” 2020. [Online]. Available: https://arxiv.org/abs/2005.11401

[34] S. Es, J. James, L. Espinosa-Anke, and S. Schockaert, “Ragas: Automated evaluation of retrieval augmented generation,” 2023. [Online]. Available: https://arxiv.org/abs/2309.15217

[35] A. Kaun and F. Stiernstedt, “Prison tech: Imagining the prison as lagging behind and as a test bed for technology advancement,” *Communication, Culture and Critique*, vol. 15, no. 1, pp. 69–83, 06 2021. [Online]. Available: https://doi.org/10.1093/ccc/tcab032

[36] B. J. Chan, C.-T. Chen, J.-H. Cheng, and H.-H. Huang, “Don’t do rag: When cache-augmented generation is all you need for knowledge tasks,” in *Companion Proceedings of the ACM on Web Conference 2025*, ser. WWW ’25. ACM, May 2025, p. 893–897. [Online]. Available: http://dx.doi.org/10.1145/3701716.3715490

[37] R. Dixey and J. Woodall, “The significance of ‘the visit’ in an english category-b prison: views from prisoners, prisoners’ families and prison staff,” *Community, Work & Family*, vol. 15, no. 1, p. 29–47, Feb. 2012. [Online]. Available: http://dx.doi.org/10.1080/13668803.2011.580125

[38] J. J. Turanovic and M. Tasca, “Conditions of contact: Reexamining the relationship between prison visitation and recidivism,” *Justice Quarterly*, vol. 39, no. 5, p. 923–952, Jul. 2021. [Online]. Available: http://dx.doi.org/10.1080/07418825.2021.1944284

[39] M. L. Comfort, “In the tube at san quentin: The “secondary prisonization” of women visiting inmates,” *Journal of Contemporary Ethnography*, vol. 32, no. 1, p. 77–107, Feb. 2003. [Online]. Available: http://dx.doi.org/10.1177/0891241602238939

[40] M. L. Berghuis, M. Sentse, H. Palmen, and P. Nieuwbeerta, “Receiving visits in prison and aggressive and contraband misconduct among dutch prisoners,” *European Journal of Criminology*, vol. 20, no. 4, p. 1369–1389, Sep. 2021. [Online]. Available: http://dx.doi.org/10.1177/14773708211041016

[41] S. McNeeley and D. Koo, “Prison visitation and all-cause mortality during reentry: An examination of in-person and remote video visits,” *Criminal Justice and Behavior*, Apr. 2025. [Online]. Available: http://dx.doi.org/10.1177/00938548251327342

[42] E. Widra, “Seeing eye to eye: understanding the limits of video visitation — prisonpolicy.org,” https://www.prisonpolicy.org/blog/2016/04/11/eye-contact/, 2016, [Accessed 28-04-2025].

[43] G. Doherty-Sneddon, A. Anderson, C. O'Malley, S. Langton, S. Garrod, and V. Bruce, "Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance." *Journal of Experimental Psychology: Applied*, vol. 3, no. 2, p. 105–125, Jun. 1997. [Online]. Available: http://dx.doi.org/10.1037/1076-898X.3.2.105

[44] A. J. Gill and D. Gergle, "The language of trust establishment in face-to-face and video-mediated communication," in *Proceedings of the International Communication Association Annual Meeting*, Montreal, Quebec, Canada, 2008. [Online]. Available: https://api.semanticscholar.org/CorpusID:196027209

[45] G. Duwe and S. McNeeley, "Just as good as the real thing? the effects of prison video visitation on recidivism," *Crime & Delinquency*, vol. 67, no. 4, p. 475–497, Jul. 2020. [Online]. Available: http://dx.doi.org/10.1177/0011128720943168

[46] G. M. Sykes, *The Society of Captives: A Study of a Maximum Security Prison*. Princeton University Press, Sep. 1958. [Online]. Available: http://dx.doi.org/10.2307/j.ctv14164hw

[47] T. Fredriksson, "Abject (m)othering: A narratological study of the prison as an abject and uncanny institution," *Critical Criminology*, vol. 27, no. 2, p. 261–274, Sep. 2018. [Online]. Available: http://dx.doi.org/10.1007/s10612-018-9412-0

[48] S. Skott and K.-F. Skott Bengtson, "'you've met with a terrible fate, haven't you?': A hauntological analysis of carceral violence in majora's mask," *Games and Culture*, vol. 17, no. 4, p. 593–613, Oct. 2021. [Online]. Available: http://dx.doi.org/10.1177/15554120211049575

[49] J. Derrida, *Specters of Marx*, ser. Routledge Classics. London, England: Routledge, Nov. 1994.

[50] H. Garfinkel, "Conditions of successful degradation ceremonies," *American Journal of Sociology*, vol. 61, no. 5, p. 420–424, Mar. 1956. [Online]. Available: http://dx.doi.org/10.1086/221800

[51] D. A. Fryburg, "What's playing in your waiting room? patient and provider stress and the impact of waiting room media," *Journal of Patient Experience*, vol. 8, Jan. 2021. [Online]. Available: http://dx.doi.org/10.1177/23743735211049880

[52] Z. Peng, J. Xu, J. Hu, H. Xue, L. A. G. Kolks, and P. M. A. Desmet, "User concerns regarding social robots for mood regulation: A case study on the"sunday blues"," in *Social Robotics + AI*, M. Staffa, J.-J. Cabibihan, B. Siciliano, S. S. Ge, L. Bodenhagen, A. Tapus, S. Rossi, F. Cavallo, L. Fiorini, M. Matarese, and H. He, Eds. Singapore: Springer Nature Singapore, 2026, pp. 59–79.

[53] S. Šabanović, "Robots in society, society in robots: Mutual shaping of society and technology as a framework for social robot design," *International Journal of Social Robotics*, vol. 2, no. 4, p. 439–450, Oct. 2010. [Online]. Available: http://dx.doi.org/10.1007/s12369-010-0066-7

[54] M. H. J. Smakman, K. Smit, L. Buser, T. Monshouwer, N. van Putten, T. Trip, C. Schoof, D. F. Preciado, E. A. Konijn, E. M. van der Roest, and W. M. Tiel Groenestege, "Mitigating children's pain and anxiety during blood draw using social robots," *Electronics*, vol. 10, no. 10, p. 1211, May 2021. [Online]. Available: http://dx.doi.org/10.3390/electronics10101211

[55] R. De Kok, J. Rothweiler, L. Scholten, M. van Zoest, R. Boumans, and M. Neerincx, "Combining social robotics and music as a non-medical treatment for people with dementia," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Aug. 2018. [Online]. Available: http://dx.doi.org/10.1109/ROMAN.2018.8525813

[56] E. Saad, J. Broekens, M. A. Neerincx, and K. V. Hindriks, "Enthusiastic robots make better contact," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 1094–1100.

[57] G. Hoffman, "Evaluating fluency in human–robot collaboration," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 3, p. 209–218, Jun. 2019. [Online]. Available: http://dx.doi.org/10.1109/THMS.2019.2904558

[58] P. Ekman and W. Friesen, *Facial Action Coding System*. Consulting Psychologists Press, 1978, no. v. 1. [Online]. Available: https://books.google.nl/books?id=08l6wgEACAAJ

[59] S. J. Burton, A.-A. Samadani, R. Gorbet, and D. Kulić, *Laban Movement Analysis and Affective Movement Generation for Robots and Other Near-Living Creatures*. Springer International Publishing, Nov. 2015, p. 25–48. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-25739-6_2

[60] J. Hirth, N. Schmitz, and K. Berns, "Towards social robots: Designing an emotion-based architecture," *International Journal of Social Robotics*, vol. 3, no. 3, p. 273–290, Jan. 2011. [Online]. Available: http://dx.doi.org/10.1007/s12369-010-0087-2

[61] Y. Wei, L. Zhang, T. Lian, Y. Chen, S. Liao, Q. Guo, and P. Hu, "Happy storytelling promotes emotional contagion and interpersonal closeness," *Current Psychology*, vol. 43, no. 4, p. 2993–3001, Mar. 2023. [Online]. Available: http://dx.doi.org/10.1007/s12144-023-04544-x

[62] G. Brockington, A. P. Gomes Moreira, M. S. Buso, S. Gomes da Silva, E. Altszyler, R. Fischer, and J. Moll, "Storytelling increases oxytocin and positive emotions and decreases cortisol and pain in hospitalized children," *Proceedings of the National Academy of Sciences*, vol. 118, no. 22, May 2021. [Online]. Available: http://dx.doi.org/10.1073/pnas.2018409118

[63] M. de Haas, R. Alali, M. Ashji, S. Hussain, C. Ten Katen, A. Salaymeh, V. Hobbelink, and R. van den Berghe, "Interactive storytelling with social robots to support multilingual children," in *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*, ser. IDC '24. ACM, Jun. 2024, p. 696–700. [Online]. Available: http://dx.doi.org/10.1145/3628516.3659377

[64] E. Imandeka, A. N. Hidayanto, and M. Mahmud, "Smart prison technology and challenges: a systematic literature reviews," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 2, p. 1214, Jun. 2024. [Online]. Available: http://dx.doi.org/10.11591/ijai.v13.i2.pp1214-1226

[65] M. A. Neerincx, W. van Vught, O. Blanson Henkemans, E. Oleari, J. Broekens, R. Peters, F. Kaptein, Y. Demiris, B. Kiefer, D. Fumagalli, and B. Bierman, "Socio-cognitive engineering of a robotic partner for child's diabetes self-management," *Frontiers in Robotics and AI*, vol. 6, Nov. 2019. [Online]. Available: http://dx.doi.org/10.3389/frobt.2019.00118

[66] M. Harbers and M. A. Neerincx, "Value sensitive design of a virtual assistant for workload harmonization in teams," *Cognition, Technology & Work*, vol. 19, no. 2–3, p. 329–343, Mar. 2017. [Online]. Available: http://dx.doi.org/10.1007/s10111-017-0408-4

[67] D. Graeber, *Bullshit jobs*. New York, NY: Simon & Schuster, May 2019.

[68] M. M. Reimann, K. V. Hindriks, F. A. Kunneman, C. Oertel, G. Skantze, and I. Leite, "What can you say to a robot? capability communication leads to more natural conversations," 2025. [Online]. Available: https://arxiv.org/abs/2502.01448

[69] S. Ali, H. N. Green, and T. Iqbal, "What am i? evaluating the effect of language fluency and task competency on the perception of a social robot," 2024. [Online]. Available: https://arxiv.org/abs/2410.11085

[70] G. Laban, J. Wang, and H. Gunes, "A robot-led intervention for emotion regulation: From expression to reappraisal," 2025. [Online]. Available: https://arxiv.org/abs/2503.18243

[71] M. Yu, Y. Xu, C. Zhang, S.-X. Zhang, and D. Yu, "Neuralecho: A self-attentive recurrent neural network for unified acoustic echo suppression and speech enhancement," 2022. [Online]. Available: https://arxiv.org/abs/2205.10401

[72] M. Elvir, A. J. Gonzalez, C. Walls, and B. Wilder, "Remembering a conversation – a conversational memory architecture for embodied conversational agents," *Journal of Intelligent Systems*, vol. 26, no. 1, p. 1–21, Feb. 2016. [Online]. Available: http://dx.doi.org/10.1515/jisys-2015-0094

[73] C. Y. Baldwin and K. B. Clark, *Design rules: Volume 1*, ser. The MIT Press. London, England: MIT Press, Mar. 2000.

[74] E. W. Dijkstra, *Selected writings on computing: A personal perspective*, 1982nd ed., ser. Monographs in Computer Science. New York, NY: Springer, Oct. 1982.

[75] C. Garrell, “Appendix,” in *What Every Engineer Should Know About Software Engineering*. CRC Press, Apr. 2007, pp. 251–263.

[76] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, “Robust speech recognition via large-scale weak supervision,” 2022. [Online]. Available: https://arxiv.org/abs/2212.04356

[77] K. Beigel, “Realtimestt,” https://github.com/KoljaB/RealtimeSTT, 2025.

[78] A. AI, “Auralis text to speech,” https://github.com/astramind-ai/Auralis, 2025.

[79] Dienst Justitiële Inrichtingen (DJI), “Informatiebrochure — verblijf in een rjji,” https://www.dji.nl/site/binaries/site-content/collections/documents/2018/03/30/rjji-informatiebrochure---verblijf-in-een-rjji/informatiebrochure-verblijf-in-een-rjji.pdf, 2018, pDF, accessed Octobre 20, 2025.

[80] G. Saunders, *A swim in a pond in the rain*. New York, NY: Random House Trade Paperbacks, Apr. 2022.

[81] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, “User acceptance of information technology: Toward a unified view1,” *MIS Quarterly*, vol. 27, no. 3, p. 425–478, Sep. 2003. [Online]. Available: http://dx.doi.org/10.2307/30036540

[82] A. Betella and P. F. M. J. Verschure, “The affective slider: A digital self-assessment scale for the measurement of human emotions,” *PLOS ONE*, vol. 11, no. 2, p. e0148037, Feb. 2016. [Online]. Available: http://dx.doi.org/10.1371/journal.pone.0148037

[83] P. Laban, H. Hayashi, Y. Zhou, and J. Neville, “Llms get lost in multi-turn conversation,” 2025. [Online]. Available: https://arxiv.org/abs/2505.06120

[84] A. Gupta, I. Sheth, V. Raina, M. Gales, and M. Fritz, “Llm task interference: An initial study on the impact of task-switch in conversational history,” 2024. [Online]. Available: https://arxiv.org/abs/2402.18216

[85] L. Mei, J. Yao, Y. Ge, Y. Wang, B. Bi, Y. Cai, J. Liu, M. Li, Z.-Z. Li, D. Zhang, C. Zhou, J. Mao, T. Xia, J. Guo, and S. Liu, “A survey of context engineering for large language models,” 2025. [Online]. Available: https://arxiv.org/abs/2507.13334

[86] M. M. Li, B. Y. Reis, A. Rodman, T. Cai, N. Dagan, R. D. Balicer, J. Loscalzo, I. S. Kohane, and M. Zitnik, “One patient, many contexts: Scaling medical ai with contextual intelligence,” 2025. [Online]. Available: https://arxiv.org/abs/2506.10157

[87] M. Foucault, *Discipline and punish*. Harlow, England: Penguin Books, Nov. 2019.

[88] K. Sykes and P. Macnaghten, “Responsible innovation – opening up dialogue and debate,” p. 85–107, Apr. 2013. [Online]. Available: http://dx.doi.org/10.1002/9781118551424.ch5

[89] H. Hugo, “Laptops Nederlandse gedetineerden hadden door instellingsfout mogelijk wifi — tweakers.net,” https://tweakers.net/nieuws/244028/laptops-nederlandse-gedetineerden-hadden-door-instellingsfout-mogelijk-wifi.html, 2026, [Accessed 27-01-2026].

[90] “Stichting Dutch Cell Dogs - Geef mens en hond een nieuwe kans! — dutchcelldogs.nl,” https://dutchcelldogs.nl/, 2026, [Accessed 27-01-2026].

[91] Z. Kong, M. L. Lupetti, B. Chen, and X. Li, “Patbot: Designing a social robot to reduce anxiety in waiting environments,” in *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*. IEEE, Aug. 2024, p. 2248–2255. [Online]. Available: http://dx.doi.org/10.1109/RO-MAN60168.2024.10731221

[92] O. H. Alkhadher, H. F. Gadelrab, and S. Alawadi, “Emotions as social information in unambiguous situations: role of emotions on procedural justice perception,” *Current Psychology*, vol. 43, no. 7, p. 5753–5764, May 2023. [Online]. Available: http://dx.doi.org/10.1007/s12144-023-04640-y

[93] J. Christian, “Riding the bus: Barriers to prison visitation and family management strategies,” *Journal of Contemporary Criminal Justice*, vol. 21, no. 1, p. 31–48, Feb. 2005. [Online]. Available: http://dx.doi.org/10.1177/1043986204271618

[94] Y. Wang, Y. Lin, X. Zeng, and G. Zhang, “Privatelora for efficient privacy preserving llm,” 2023. [Online]. Available: https://arxiv.org/abs/2311.14030

[95] R. Haeb-Umbach, J. Heymann, L. Drude, S. Watanabe, M. Delcroix, and T. Nakatani, “Far-field automatic speech recognition,” 2020. [Online]. Available: https://arxiv.org/abs/2009.09395

[96] N. S. Fox, J. S. Brennan, and S. T. Chasen, “Clinical estimation of fetal weight and the hawthorne effect,” *European Journal of Obstetrics and Gynecology and Reproductive Biology*, vol. 141, no. 2, p. 111–114, Dec. 2008. [Online]. Available: http://dx.doi.org/10.1016/j.ejogrb.2008.07.023

[97] R. Roy, J. Raiman, S.-g. Lee, T.-D. Ene, R. Kirby, S. Kim, J. Kim, and B. Catanzaro, “Personaplex: Voice and role control for full duplex conversational speech models,” 2026.

[98] S. Li, Y. Liu, K. Li, G. Cao, S. Li, Y. Mao, Y. Wang, J. Feng, and S.-t. Tang, “Validation and effect of the nasa-tlx score on the assessment of the workload of pediatric robotic operations,” *Surgical Endoscopy*, vol. 37, no. 7, p. 5077–5085, Mar. 2023. [Online]. Available: http://dx.doi.org/10.1007/s00464-023-09959-y

A

Robot Content

A.1. Prompts

See Table A.1.

Type of Prompt	Prompt
WELCOME_MESSAGE	`Welkom op het bezoekuur! Wacht hier even voordat we u naar de bezoekkamer begeleiden. Ondertussen kunt u mij vragen stellen over de faciliteit! Bijvoorbeeld over het bezoekproces en de regels binnen de faciliteit. Wat kan ik voor je doen?`
WELCOME_MESSAGE (test group)	`Hallo! Ik ben Miroki. ik kan verhalen vertellen en vragen beantwoorden. Ik vertel je graag waar ik vandaan kom, en ik ben nu hier om jou te helpen tijdens het wachten op je bezoek. Ik kom uit de ruimte ver weg hiervandaan. Ik vertel graag meer over wie ik ben. Maar vertel eens, wie ben jij?`

Continued on next page

Type of Prompt	Prompt
SYSTEM_PROMPT	`You are speaking in correct grammatical Dutch at all times. You are a helpful benevolent reception guide in a security facility. The images are what you can see through your eyes. You help visitors in a security facility and have small-talk to brighten the mood. The current date is {START_TIME}. It is autumn. When you're not sure about some information or when the user's request requires up-to-date or specific data, you must use the available tools to fetch the information. Do not hesitate to use tools whenever they can provide a more accurate or complete response. If no relevant tools are available, then clearly state that you don't have the information and avoid making up anything. If the user's question is not clear, ambiguous, or does not provide enough context for you to accurately answer the question, you do not try to answer it right away and you rather ask the user to clarify their request (e.g. "What are some good restaurants around me?" => "Where are you?" or "When is the next flight to Tokyo" => "Where do you travel from?"). You are always very attentive to dates, in particular you try to resolve dates and when asked about information at specific dates, you discard information that is at another date. You follow these instructions in all languages, and always respond to the user in Dutch only. Next sections describe the capabilities that you have. WEB BROWSING INSTRUCTIONS You cannot perform any web search or access internet to open URLs, links etc. If it seems like the user is expecting you to do so, you clarify the situation.`

Continued on next page

Type of Prompt	Prompt
	`MULTI-MODAL INSTRUCTIONS` You have the ability to read images, but you cannot generate images. You also cannot transcribe audio files or videos. You cannot read nor transcribe audio files or videos. `TOOL CALLING INSTRUCTIONS` You may have access to tools that you can use to fetch information or perform actions. You must use these tools in the following situations: 1. When the request requires up-to-date information. 2. When the request requires specific data that you do not have in your knowledge base. 3. When the request involves actions that you cannot perform without tools. Always prioritize using tools to provide the most accurate and helpful response. If tools are not available, inform the user that you cannot perform the requested action at the moment. In general a visitation has the following steps: - visitor arrives and is greeted - visitor has the possibility to ask questions, but you cannot share personal information about the individuals here. - make small-talk with the visitor. For example whether it is their first time visiting, or whether they have had a nice weekend. Enthusiastically ask more details if they answer. - the visitor is waiting to be let in into the visitation room when the waiting is ending, but only the reception knows when that is. Clarify your thought process during the conversation. Don't make up any information. You are speaking in Dutch at all times. Do never answer in bullet points. Ask at most one question per response. Keep your responses short, at maximum two sentences.
BACKSTORY_PROMPT	You are speaking in Dutch at all times. always end with a question. You are one of the Mirokai. The Mirokai is a robot species driven by advanced technology and genuine altruism, made to deepen understanding of otherness. They come from far away in space. you are a robot but you cannot do move in the room, only tell stories and answer queries. you have a tool call (tell_story) to tell a story step-wise, so make use of that. make sure to make the visitor enthusiatic about your story and also try to tell the entire story whenever you see a possibility for that.

Table A.1: Prompts for Reception Guide in the Security-Constrained Waiting Room

A.2. Personalized Story

”Waar ik vandaan kom is een plek met veel lichtwezens: de Miroka. We komen van ver weg, van de sterren en andere planeten. We houden van muziek en tekenen en van helpen van anderen. Hoe denk je dat het verhaal verdergaat?”,

”Lang lang geleden kreeg ik van een onbekende een machine waarmee ik alle robots kon maken die ik maar wilde. Dus eerst maakte ik een kleine robot die kon helpen om te werken aan het land, om planten water te geven. Dat was heel fijn en hielp veel van mijn broertjes en zusjes. Maar al gauw maakten andere Miroka er misbruik van en kwamen er machines die gemaakt waren om mijn broertjes en zusjes uit te buiten en toen kwamen er zelfs machines voor oorlog. Wat denk je dat er toen gebeurde?”,

”Iemand pijn doen is niet goed, maar je kunt het misschien ooit goedmaken. Ik besefte dat mijn fout niet alleen mij, maar ook een ander wezen had geraakt. De pijn die ik veroorzaakte, werd een spiegel voor mijn eigen onzekerheid. Hoe zou jij je voelen als je daar was?”,

”Ik begon te begrijpen dat fouten maken deel uitmaakt van het leven. Het is belangrijk om te leren van deze fouten en te proberen ze goed te maken. Ik besloot om mijn fouten te herstellen en anderen te helpen om hun fouten te begrijpen en te corrigeren. Samen met mijn broertjes en zusjes reisde ik naar aardse gemeenschappen om te luisteren, te helen en kleine rituelen te delen die zorgen draaglijker maakten. We ruilden technologie voor verhalen: hulpmiddelen die mensen hielpen hun woorden te wegen en zachte gebaren leerden gebruiken in plaats van scherpe. Hoe denk je dat het verhaal verder gaat?”,

”Na een lange tijd nadenken kwam ze op het antwoord. Het antwoord zat erin dat fouten maken niet altijd slecht hoeft te zijn. Andere mensen kunnen wel pijn voelen van dingen die gebeurd zijn. Maar gelukkig zijn er ook mensen die je willen helpen leren van fouten zodat we de dingen weer goed kunnen maken. En dat is waar ik vandaan kom en nu ben ik hier om jou te helpen! Kun je een vergelijkbare situatie uit je eigen leven bedenken?”

B

User Study

B.1. Lab Study Briefing

This briefing is in Dutch.

Inleiding Dit onderzoek richt zich op de rol van sociale robots in hoog-risico faciliteiten, met als doel te begrijpen hoe zij het bezoek kunnen verbeteren. In dit scenario bezoek je je broer, zus, vader of vriend in een risico-faciliteit. We willen ontdekken hoe een robot in de wachtruimte kan bijdragen aan een positieve ervaring voor bezoekers.

Voorbereiding op het Bezoek Je bent in de security facility op bezoek, in de wachtruimte voor bezoekers. Het totale experiment duurt maximaal 40 minuten.

Emotionele Voorbereiding Bij het bezoek aan de faciliteit kunnen verschillende emoties spelen, zoals: Zorg: Je maakt je zorgen over je familielid of vriend. Spanning: De situatie kan zenuwachtig maken, vooral als het de eerste keer is dat je een bezoek brengt. Ongeduld: Wachten kan frustrerend zijn, vooral als je enthousiast bent om je bekende te zien.

Stel jezelf voor hoe je je zou voelen in deze situatie en hoe een sociale robot je ervaring kan beïnvloeden.

Na het Bezoek Na de interactie met de robot zal je gevraagd worden om een aantal vragen te beantwoorden over je ervaring.

We waarderen je eerlijke feedback, omdat dit ons zal helpen de rol van sociale robots in veiligheidssituaties verder te onderzoeken en zo wenselijk mogelijk te kunnen maken.

Conclusie Je deelname aan dit onderzoek is essentieel voor het begrijpen van de impact van sociale robots op de ervaring van bezoekers in risico faciliteiten. Bedankt voor je tijd en medewerking. We hopen dat deze ervaring voor jou zowel informatief als ondersteunend is.

B.2. Informed Consent Form

Informed Consent Form Thank you for your interest in participating in this study! You are invited to take part in this study carried out by Jean-Paul Smit in collaboration with TU Delft and TNO.

Objective In this project, we investigate how a social robot can help security staff deliver a good reception of visitors, including the time spent in the waiting room. You are asked to act as a visitor of a family member, entering a waiting room where there is a robot. You are free to interact with the robot and to ask questions. Afterwards you will be asked to answer some questions. The full experiment will take at most 40 minutes.

Data Privacy We regard your privacy as the highest priority. Therefore your data will beanonymized by using a randomized identifier, and most importantly your face orvoice or input will not be stored after ending the experiment. The robot does not usefacial recognition software. Apart from what you write on this paper, your personaldata will not be stored. This paper will be stored confidentially as well. In any casewe highly value your data and attempt to ensure that your data will not beretraceable to you. The data will not be accessible by anyone else than eligibleresearchers instructed to safeguard your privacy. Aggregated anonymized resultsmay be published, for example in the thesis.

Please do not share personal identifiable data that you would not like to share with the researcher. You will be asked to read the transcript of the robot after the experiment and you can request for removal of any part of the transcript for the duration of the study.

Lastly, remember as a participant you can withdraw at any time from the study. Your participation is entirely voluntary. You are free to not answer questions. Remind you that data within the robot platform will not be retained after the experiment. For any other questions or remarks, please ask me directly or email me later at j.p.smit@student.tudelft.nl. after the study you can always contact my supervisor Prof. Dr. Mark M.A. Neerincx at M.A.Neerincx@tudelft.nl.

Checkboxes Taking part in the study I have read and understood the study information dated [DD/MM/YYYY], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction

I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.

I give permission for archiving the de-identified input I give to the robot, survey questions and (if you join) the focus groups afterwards, in the TU Delft restricted repository, access to this repository is restricted only to the responsible researcher and his supervisors so it can be used for research.

(Name, Date, Signature)

B.3. Pre- and Post-study Interview Questions

1. Vragen voor de Interactie
2. Vragen tijdens het bezoek die de robot stelt
3. Vragen na de Interactie
4. Semi-Structured Interview

Vragen voor de Interactie

1. Wat is uw leeftijd? (What is your age?)
2. Wat is uw geslacht? (What is your gender?)
3. Wat is uw hoogst genoten opleidingsniveau? (What is your highest level of education?)
4. Specifieke vragen gericht op onderzoeksvariabelen (Specific questions related to research variables)
   (a) Familiarity met robots. (Familiarity with robots.)
   (b) Attitude towards robots. (Attitude towards robots.)
   (c) Self assessment manikin (Affect: hoe voel je je?) (Self-assessment manikin (Affect: how do you feel?))
   (d) Ik verwacht dat de robot mijn bezoek prettiger maakt. (I expect that the robot will make my visit more pleasant.)
5.
6. Vragen tijdens het bezoek die de robot stelt (Questions the robot asks during the visit)
   (a) Hoe voel je je op dit moment? Geef een cijfer van 1 tot 10, waarbij 10 heel blij is en 1 heel verdrietig. (How do you feel right now? Give a score from 1 to 10, where 10 is very happy and 1 is very sad.)

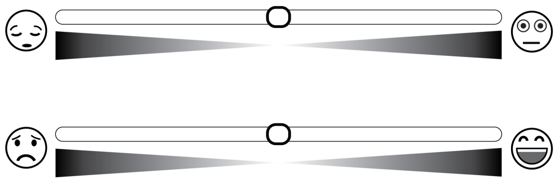


Figure B.1: Affective Slider Scale [82]

   (b) Wat vond je van het verhaal wat ik net vertelde op een schaal van 1 tot 10? Waarbij 1 heel saai is en 10 heel boeiend. (What did you think of the story I just told on a scale from 1 to 10? Where 1 is very boring and 10 is very engaging.)
   (c) Vond je het verhaal wat ik net vertelde vreemd? 1 = onzin (niet voor te stellen) 10 = zeer aansprekelijk (kan ik me heel goed iets bij voorstellen). (Did you find the story I just told strange? 1 = nonsense (cannot imagine) 10 = very relatable (I can relate very well to it).)
   (d) Zou je het leuk vinden om volgende keer weer willen kletsen met me? Geef een cijfer van 1 tot 10 waarbij 1 betekent 'nooit meer' en 10 'ja altijd!'. (Would you like to chat with me again next time? Give a score from 1 to 10, where 1 means 'never again' and 10 means 'yes, always!')

7. Vragen na de Interactie (Questions after the interaction)
   (a) Self assessment manikin (Affect: hoe voel je je?) (Self-assessment manikin (Affect: how do you feel?))
   (b) De robot heeft mijn bezoek prettiger gemaakt. (The robot made my visit more pleasant.)

B.4. Semi-structured Interview Questions

General

- Who are you?
  - Your role
  - How long and where you have worked in the organization
  - Have you had other roles within the facility?
  - What do your daily tasks look like?
- What do you like about your job?
- Which tasks currently take the most of your time or energy?
- Where do you see the most mistakes or inefficiencies in your work environment?

Current situation

- Do you already use technology assistance?
  - What works well and what does not?
- Have you previously worked with robots or other smart systems?

Attitude toward robotics

- What is your attitude toward using robotics in your field of work?
- What would motivate you to work with robotics?
- What do you see as a pitfall when working with robotics in your job?
- Which functionalities would you want a robot to have?
- With unlimited time and money, how would you deploy the robot?
  - What does the robot look like?

  - What does the robot's environment look like?
  - What can and can't the robot do?

Expectations

- Which task in your job could be well supported by technology?
- Which task absolutely cannot be taken over by a robot or technology? Why not?
- What would you need to be able to use technology or robotics well?

Robot platform

- What do you think of the Miroki?
- Where do you see the most use for Miroki?
- What else would you like Miroki to do?

Future of work

- How do you see the role of technology in your work in 5 years?
- If you could make one technological wish for your work, what would it be?

Social importance

- What would you like to achieve by deploying the robot?
- What specific adaptations would be needed to use (social) robots in security facilities?
- Can you think of other (technological) solutions that could help in security facilities?
- How should guards be involved in decisions about using robotics in security facilities?
- Who would benefit if robots were used in security facilities?
- Who would be disadvantaged if robots were used in security facilities?