



HAL
open science

A Human-Robot Mutual Learning System with Affect-Grounded Language Acquisition and Differential Outcomes Training

Alva Markelius, Sofia Sjöberg, Zakaria Lemhauori, Laura Cohen, Martin Bergström, Robert Lowe, Lola Cañamero

► **To cite this version:**

Alva Markelius, Sofia Sjöberg, Zakaria Lemhauori, Laura Cohen, Martin Bergström, et al.. A Human-Robot Mutual Learning System with Affect-Grounded Language Acquisition and Differential Outcomes Training. 15th International Conference on Social Robotics (ICSR 2023), Dec 2023, Doha, Qatar. 10.1007/978-981-99-8718-4_10. hal-04483870

HAL Id: hal-04483870

<https://hal.science/hal-04483870v1>

Submitted on 29 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Human-Robot Mutual Learning System with Affect-Grounded Language Acquisition and Differential Outcomes Training ^{*}

Alva Markelius^{1,2}[0009-0003-4580-9997]^{**}, Sofia Sjöberg², Zakaria Lemhauri^{3,4}[0009-0008-0235-8897], Laura Cohen³, Martin Bergström², Robert Lowe²[0000-0002-0307-3171], and Lola Cañamero³[0000-0001-6903-1348]

¹ University of Cambridge, Department of Philosophy, Sidgwick Ave, Cambridge, UK

² DICE lab, University of Gothenburg, Department of Applied IT, Forskningsgängen 6, Gothenburg, Sweden

³ ETIS Lab, CY Cergy Paris University - ENSEA - CNRS UMR8051, France

⁴ Artificial Intelligence Lab, Vrije Universiteit Brussel (VUB), Pleinlaan 9, Brussels, Belgium

Abstract. This paper presents a novel human-robot interaction setup for robot and human learning of symbolic language for identifying robot homeostatic needs. The robot and human learn to use and respond to the same language symbols that convey homeostatic needs and the stimuli that satisfy the homeostatic needs, respectively. We adopted a differential outcomes training (DOT) protocol whereby the robot provides feedback specific (differential) to its internal needs (e.g. ‘hunger’) when satisfied by the correct stimulus (e.g. cookie). We found evidence that DOT can enhance the human’s learning efficiency, which in turn enables more efficient robot language acquisition. The robot used in the study has a vocabulary similar to that of a human infant in the linguistic “babbling” phase. The robot software architecture is built upon a model for affect-grounded language acquisition where the robot associates vocabulary with internal needs (hunger, thirst, curiosity) through interactions with the human. The paper presents the results of an initial pilot study conducted with the interactive setup, which reveal that the robot’s language acquisition achieves higher convergence rate in the DOT condition compared to the non-DOT control condition. Additionally, participants reported positive affective experiences, feeling of being in control, and an empathetic connection with the robot. This mutual learning (teacher-student learning) approach offers a potential contribution of facilitating cognitive interventions with DOT (e.g. for people with dementia) through increased therapy adherence as a result of engaging humans more in training tasks by taking an active teaching-learning role. The homeostatic motivational grounding of the robot’s language acquisition has potential to contribute to more ecologically valid and social (collaborative/nurturing) interactions with robots.

Keywords: Language Acquisition · Developmental Robotics · Socially Assistive Robotics · Mutual Learning · Human-Robot Collaboration

^{*} All authors declare that they have no conflicts of interest. Supported by an EUTOPIA Undergraduate Research Support Scheme (EURSS) grant to AM and SS for a research stay at ETIS, and by Brain +. The work was carried out partly at ETIS, CY Cergy Paris University, and partly at the Dept. of Applied IT, University of Gothenburg.

^{**} Corresponding author’s email: ajkm4@cam.ac.uk

1 Introduction

The increased implementation of socially assistive robots (SAR) to provide motivational, social, pedagogical, and therapeutic assistance raises the questions of how components such as emotion expression, language and motivation can be grounded in homeostatic interaction and collaboration with environment and other social agents [25, 28]. This is particularly relevant in order for SAR to act appropriately and to enable e.g. formation of affective social bonds [17], adaptive acting upon and expressing emotion [25] and behaviour grounded in motivation, physiological needs and environmental conditions [9, 21]. These components are important to facilitate interactions e.g., for vulnerable populations, which constitute one of the main target groups for SAR [28], including elderly and children in the contexts of learning and cognitive interventions. Human-robot collaboration and interaction setups with SAR are gaining traction as a promising approach for facilitating and mitigating cognitive interventions [2]. This includes memory training for people with e.g. mild cognitive impairment (MCI), which is a clinical phase between normal aging and dementia [15] or dementia [1, 4, 18]. Cognitive interventions have emerged as crucial alternatives to pharmacological treatments as they provide accessible, effective, and scalable solutions to address issues that often have limited or no available pharmacological alternatives [3]. In particular, for patients with MCI, cognitive training is gaining attention as a promising form of intervention because of its effectiveness in slowing down, or even halting the advance to dementia [16]. However, treatment adherence in cognitive interventions tends to be low, in particular for older adults, resulting in discontinued or uncompleted treatments [14, 36]. This issue is likely the result of lack of engagement in the treatments, both on an affective level and as a result of boredom, lack of social accountability and means to consistently carry out the intervention.

Previous research has included SAR to facilitate engagement in cognitive interventions, where usually the robot takes a supporting and teaching role [18]. We present a collaborative human-robot interactive setup that includes a mutual learning system for both robot and human, offering a novel contribution to implementing SAR for cognitive interventions. Mutual learning has great potential to facilitate both the advancement of developmental robotics as well as memory training in humans. We define it here as teacher-student collaborative learning, where both robot and human act as teacher and student simultaneously, such as in the learning paradigm described in [13]. The setup allows the human to take a more active role as a caregiver to a robot learning to acquire language through homeostatic social and affective interaction. As such, this approach is expected to foster greater engagement, a heightened sense of control, and a deeper feeling of contribution than previous approaches where the human has a more passive role. Additionally, the present study includes differential outcomes training (DOT), as a methodology to enhance learning for the human and subsequently increase robot learning efficiency (as described in the following sections). This paper accounts for the methodology of the present collaborative setup and presents the results from an initial pilot study conducted to gain insights into the social and affective experience of the humans and its effectiveness for the robot software architecture in this specific context. This is the first time to the authors knowledge that a human-robot mutual learning methodology has been developed that includes specialised training protocols, e.g. DOT for the human and homeostatic language acquisition for the robot in combination to facilitate the learning of both. Additionally, as DOT is a training protocol specifically suited for neurodegenerative disorders [32, 35], this methodology offers a novel way for the human to be more proactive in the training, as compared to many other cognitive training interventions. The rest of the paper is set out as follows: Section 2 accounts for previous related research; Section 3 provides an overview of the methodological approach in the present study; Sections 4 and 5 account for pilot study results and discussion, respectively.

2 Related work

The work presented in this paper lies in the intersection of developmental robotics (including language acquisition, motivation and affect development) and research into the use of homeostatic mechanisms in robotics (including emotion and social grounding of interactive robot activity). Finally, one of the main components of this work includes memory training of the human part of the interactive scenario, which

builds on previous work in cognitive science and psychology (differential outcomes training and cognitive interventions).

2.1 Homeostatic Developmental Robotics and Language

This study uses and builds upon a previously developed human/robot interaction setup developed by Lemhaouri et.al. [20]. The original setup concerns language acquisition of a humanoid robot through interaction with a human caregiver and is grounded in affective motivation as the main driving force behind language acquisition. In this setup, the robot possesses a vocabulary similar to that of an 8-month-old infant during the linguistic babbling phase [10]. The setup is based on trial-and-error, where the robot expresses an internal need and the caregiver has to learn this need by probing, i.e. providing the robot with an object to fulfill the need. If guessed correctly, the word (expressed need) will be reinforced to be associated with the particular need. Previous homeostasis based work in robotics has focused on biologically inspired approaches to facilitate interaction between robots and humans and robot inner motivations as a driving factor to foster decision making, behavior and emotion in relation to environment and other social agents [9, 37]. For instance, dynamic, biologically-inspired social mechanisms in artificial agents have been investigated for forming long-term affective bonds [17] and social learning [6] as well as more complex mechanisms such as reasoning and interacting on perceived affective and intentional state [31]. This has previously been shown to be relevant to robots deployed in elderly care [37] where the inner motivations of the robot enable it to assist, converse and care depending on environmental and social stimuli. This functional approach differs from many other computational approaches of robot language development as it acknowledges the role of motivation, affect and semantic meaning in linguistic acquisition [20]. Section 3.1 provides a detailed overview of the methodology and the robot software architecture.

2.2 Differential Outcomes Training

DOT refers to the learning of unique stimulus-response pairs when certain stimuli are associated with certain rewards/outcomes. More specifically, DOT entails correct responses on a given task being reinforced by feedback that is unique to both the response and stimulus. For example, the response of pushing forward a "cookie" object (response) following the stimulus (utterance "hungry" in an unknown language) would yield the rewarding outcome of a specific social expression (e.g. smile). For different stimulus-response pairings (with a correct response) another unique outcome would be expressed (e.g. head nod). In non-differential outcomes tasks, this feedback following the correct response would be randomized (either smile or head nod irrespective of stimulus-response). The DOT approach, contrasting with standard reward-based training where the same reward is given irrespective of stimulus and response 'pairings', has consistently been shown to enhance discriminative learning and memory performance accuracy in animals, children and adults [11, 12, 30]. Furthermore, evidence suggests it is effective in elderly people with memory impairments such as Alzheimer's disease [32, 35]. A recent systematic review with meta-analysis [29] revealed that the impact of DOT on learning and memory exhibits a medium to large effect size as compared to control condition comparisons. This review also found that the positive learning effect of DOT was found in several studies deploying persons with dementia or MCI. The use of repeated sessions of training (i.e. for full cognitive interventions) may, thereby, be worth exploring as a means for mitigating the symptoms of early stage dementia or MCI. Our previous work has suggested using DOT in combination with SAR [26] with promising results, showing that the DOT effect can also be obtained in the context of human-robot interaction. DOT-effects have been found in human-human interactive learning [33] and also theorized in the context of human-robot interactive learning using facial expressions and perceived outcomes of others [23]. Furthermore, differential outcomes expectancies have been computationally modelled in the context of learning and interaction using reinforcement learning [24]. In this study, the DOT is presented in the interaction between the human and robot, as the feedback provided by the robot is differential to certain homeostatic motivations and needs (based on reinforcement learning). Furthermore, the nature of the interaction entails a teacher-student relationship, whereby the human and robot serve as both teacher and student to each other (mutual learning). Section 3.2 provides a more detailed overview of how DOT is implemented in this setup.

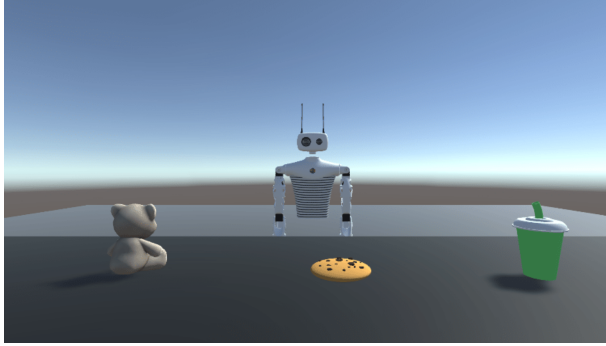


Fig. 1. The robot as seen by the participants in the study (human caregivers). The visual field includes the three objects that can be given to the robot to fulfill its internal needs

3 Methodology

The robot used in this study is the simulated (Unity implemented, version 2020.3.11f1) SDK-version of the Reachy robot (version SDK 0.5.1), developed by Pollen Robotics, a humanoid robot able to expressively move its head, antennae and arms (See Figure 1). The simulated setup closely reflects the dynamics and functions of a real world version of the setup, using the physical version of the Reachy robot (which has previously been tested and piloted). Thus includes the software architecture, the roles of the human and robot and the setup design and procedure.

3.1 Robot Software Architecture

The robot software architecture is built upon [20] who developed a model for affect grounded language acquisition, which is based on creating associations between vocabulary and internal needs through human ‘caregiver’ interaction. The setup includes a reinforcement learning cycle (using a multi-armed bandit algorithm) where the robot verbally expresses an internal need as determined by affective motivation by requesting an object from the caregiver. There are three different internal needs (hunger, thirst and curiosity) and three respective objects that can fulfill those needs (cookie, drink, teddy bear). Initially, the robot generates words randomly when one need becomes more significant than the others, the caregiver being unaware of which need the robot is expressing. Upon hearing the robot’s vocalization, the caregiver selects an object and hands it to the robot. If the provided object satisfies the robot’s need, the motivation associated with that need diminishes, and the robot receives a reward of +1, expressing its satisfaction through giving positive audiovisual feedback. However, if the object does not fulfill need, the word is penalized with a reward of -1, reducing the probability of using that word in a similar context, and the robot gives negative audiovisual feedback [20]. The robot software learning model consisted originally of three modules, the motivation, visual perception, and phonological modules [20]. In the present study, a fourth module was added, the DOT feedback module.

Motivation Module: The three internal needs (hunger, thirst, curiosity) are determined by a homeostatic variable that continuously decreases over time until the need is fulfilled, which makes it increase. This can be compared to blood glucose level that decreases over time and increases after a meal. The robot drive $d_i(t)$, representing the urge to act and satisfy the need i , is defined as the difference between the current homeostatic variable and its optimal value. The robot’s motivation to satisfy a need depends on the related drive and the intensity of the stimulus that can satisfy it [5]:

$$m_i(t) = d_i(t) + d_i(t) \cdot s_i \quad (1)$$

s_i is the intensity of the stimulus estimated by the visual perception module (described below). The robot’s motivation to act on a certain internal need is determined by a threshold that, if it is reached for a

certain need before another, leads the robot to express a ‘will’ to satisfy that need (babble) based on the phonological module. The need is satisfied based on the visual perception module (if the robot is given the right object to satisfy the need or not).

Visual Perception Module: This module helps modulate the motivation system and determines if a need is satisfied based on the robot’s ability to recognise objects given by the caregiver (cookie, drink, teddy bear) in response to its expression of an internal need. This gives the robot the ability to identify and label objects within its visual environment, while also visually associating the internal needs that the recognised objects can satisfy. The model used for object detection and recognition is a combination of a modified version of Kohonen’s self-organizing map clustering algorithm [19] and a multi-class perceptron in which the weights update $\Delta\omega$ follows the Widrow-Hoff learning rule:

$$\Delta\omega_{ij} = \epsilon VF_i(RIS_j - ISP_j) \quad (2)$$

With :

VF: The visual features of the object calculated by SIFT algorithm [22].

ϵ : the learning rate.

RIS: the robot internal state.

ISP: internal state prediction.

Phonological Module: This module allows the robot to verbally express its internal needs by a text-to-speech functionality. The robot’s language repertoire comprises two-syllable words which are chosen from some of the most frequent syllables commonly used by an 8-month-old human infant (examples include “nana”, “wada”, “pada”). The robot initially starts by randomly selecting a word whenever one particular internal need becomes more pronounced than the others. As the robot receives an object from the caregiver it learns to associate certain words in its vocabulary with certain internal needs.

DOT Feedback Module: In addition to the three above mentioned modules, the current study included a DOT feedback module. This module determines the type of positive feedback given for a certain internal need. If the robot’s need is not fulfilled the robot always expresses the same negative audiovisual feedback (looking down, lowering antenna and making a sad sound). However, if the need is fulfilled, the robot expresses a certain positive feedback associated with each internal need. Each need had different happy beeping sounds and different movements (curiosity: wagging antennae, hunger: happy arm movement, thirst: nodding head movement). In other words, each internal need has its own sound and movement associated with it, which the robot consistently expresses when that need is satisfied. Each feedback response is therefore differential with respect to each internal need fulfillment. In the non-DOT condition the feedback included the same three movements but they were instead randomised for internal needs (i.e. not tied to specific needs). For a comparison of how the interactive scenario differs between the DOT and non-DOT based feedback see Figure 2. The movements were implemented as inspired by how animals tend to express contentment or happiness (such as dogs wagging their tail). Tails and wagging motions have previously been implemented with robots to express happiness [34], and in the present study we implemented a similar wagging motion on both the robot’s antennae and arms.

3.2 Experimental Setup

The setup was tested in an initial pilot study to try out the functionality of the robot architecture including the new DOT module as well as getting some first insights into participants’ experience of interacting with the robot in this particular setup. The pilot study had only one independent variable (two levels): with or without the DOT feedback module (condition 1: DOT, condition 2: non-DOT). Furthermore, the pilot study solely included a virtual version of the robot and the interaction was thus carried out through a laptop interface. For a visualisation of the robot as seen by the participants see Fig 1. Six healthy Swedish

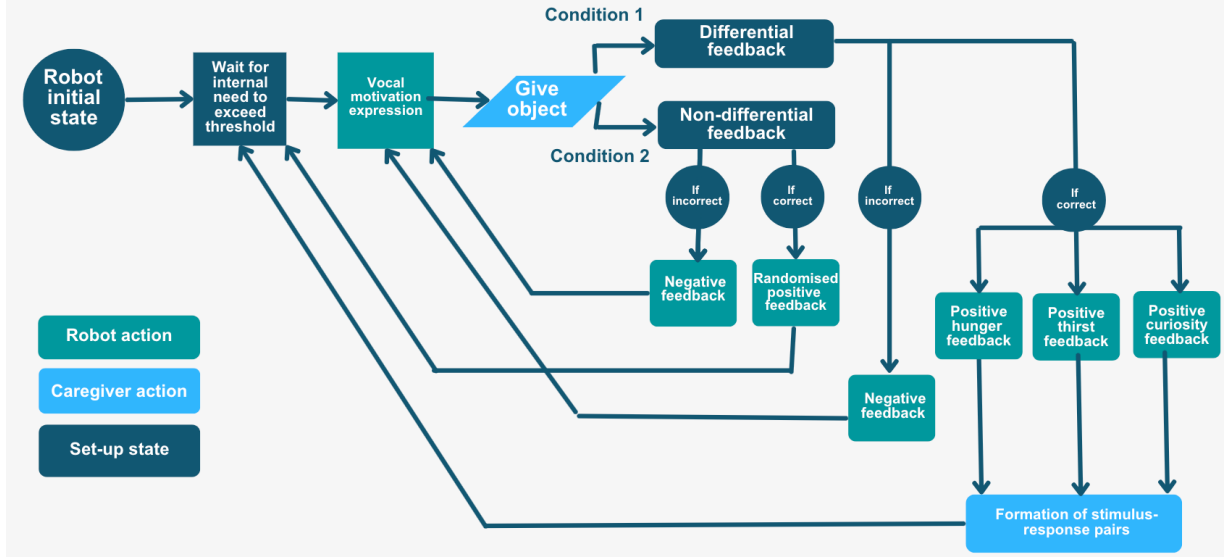


Fig. 2. Flowchart of one iteration of the experimental procedure, depending on either DOT or non-DOT conditions. The experiment consisted of 12-16 iterations depending on the fulfillment of the robot needs.

adults (not suffering from cognitive impairments or memory deficiencies) participated in the pilot study, aged between 24 and 37 years ($M=29$). Half of the participants were in the DOT condition and half in non-DOT condition. Informed consent was ensured by all of the volunteering participants by provision of informed written consent forms. The study was conducted in accordance with the WMA Declaration of Helsinki. The participants did not receive any compensation.

Procedure The experiment was carried out according to the following procedure. After having been given instructions and information about the study the participant sat down in front of the laptop with the Unity virtual version of the robot setup (as seen in Fig 1). They were instructed to give the robot an object when it expressed an internal need, and to remember what object was associated with what word. The experiment lasted for approximately 13-16 iterations which took about 10 minutes, iterations ended automatically as a function of learning. For a full flowchart of one experimental iteration see Fig 2. After the interaction with the robot, the participants filled in a survey about their experience and affective state associated with the interaction. Finally, a semi-structured interview was conducted with each participant about their experience and perception of the interactive scenario and the robot.

Data Collection The robot's ability to learn to associate certain words in its vocabulary with a certain internal need is determined by convergence rate. Full convergence is reached when the robot achieves a state where it consistently selects appropriate words based on its internal state. Additionally, this enables the caregiver to understand the robot, facilitating the robot's ability to obtain the desired objects. The evaluation metric used to determine convergence rate is the moving average of rewards (MAR) received by the robot:

$$MAR_n = \begin{cases} \frac{1}{n} \sum_{i=1}^n R_i & \text{if } n < m \\ \frac{1}{m} \sum_{i=n-m+1}^n R_i & \text{Otherwise} \end{cases} \quad (3)$$

m is the number of previous reward values used to calculate the MAR at the iteration n . In this experiment we chose $m=5$. The convergence time, on the other hand, is defined as the number of iterations required for

the robot to reach 90% convergence. Thus, the dependent variable in terms of robot language acquisition is average convergence rate. For determining the participants' affective experience of the setup self-report data was collected using the dimensional Self Assessment Manikin (SAM) Scale [7]. It was used with 3 dimensions, valence (happy/unhappy), arousal (stressed/not stressed) and dominance (in control/not in control) each to be rated between 1 to 5. To complement the scale, the survey also contained likert-scale based questions about the specific setup. Finally, the semi-structured interviews were recorded and transcribed to enable a thematic analysis as well as case by case qualitative analysis. The thematic analysis was conducted by two of the authors individually to encompass a more comprehensive and objective identification of themes and key words. Initially, general code words were identified in the transcribed interviews, and were then categorised into more general themes. The most commonly occurring themes, mostly relevant to the objectives of the study were then used as data for the analysis.

4 Results

This section provides the results of the pilot study conducted to constitute an initial test of the collaborative interaction setup. The first section gives an overview of robot language acquisition. The second section presents the results of the participants' experience as obtained from the survey and semi-structured interview thematic analysis.

4.1 Robot Language Acquisition

The robot's ability to learn to associate certain words in its vocabulary with a certain internal need is determined by convergence rate (as described in 2.2). Figure 3 visualises the mean convergence rate for DOT (left) and non-DOT (right) conditions, respectively. No condition met full convergence, possibly due to the limited number of iterations. However, in the DOT condition a total higher average reward was achieved than in the non-DOT condition.

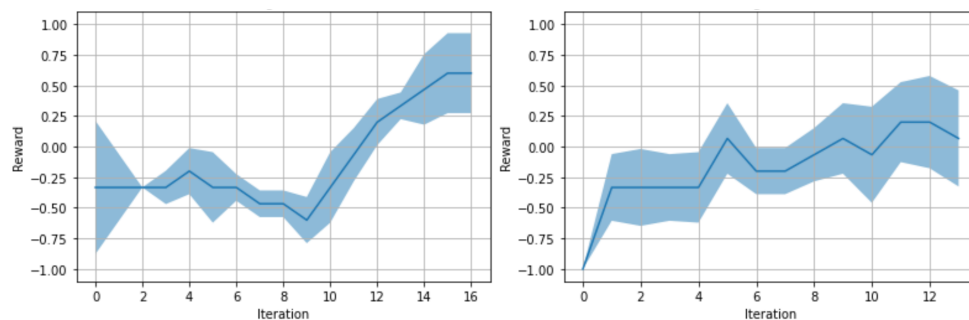


Fig. 3. Average reward and standard deviation obtained by the robot in the DOT (left) and non-DOT (right) condition over the mean number of iterations. Iterations ended automatically as a function of learning.

4.2 Participant Experience

The mean self-reported affective ratings of the SAM-scale are displayed in Figure 4. Participants generally reported high valence (happy) ($M=3.5$) and dominance (feeling of being in control) ($M=3.7$) as well as low arousal (not stressed) ($M=2.5$). The results of the additional questions about the participants' experiences can be seen in Table 1.

Furthermore, a thematic analysis of the semi-structured interviews revealed some insights into the participants' experience of interacting with the robot and the setup. Most participants found the task to be easy to understand, although lacking in being grounded in a sufficient purpose. Most participants

found it easy to differentiate whether the robot was given an object it wanted or not. Commonly occurring themes were **confusion** but at the same time **amusement** and **excitement**. Another theme was **negative/positive coding** of the robot’s movements and sounds (as expressed as audiovisual feedback); most participants found the antennae to be natural and intuitive to interpret as positive or negative. However, a commonly occurring theme regarding the robot’s arm movement was **uncanniness/unnaturalness**. There also seemed to be a general confusion about whether certain reactions were stronger than others (e.g. if arm movements meant *more* happy than e.g. head movements). This was expressed by a participant saying “*The logic of the antennas made a lot of sense. But I did not understand the intensity of the body language, if certain movements were more positive than others.*” Another participant stated that “*...it has too large arms and it was unsettling, but the rest of the movements were pretty cute. The antennae felt like ears and since animals are familiar when they go up or down for happy and sad the antennae worked the same way.*” Another common theme concerned the robot’s **age** and **child-likeness**. Most participants expressed some form of care or empathy towards the robot, which was generally perceived as a 1-3 year old, evoking feelings of wanting to provide it with its needs. One participant reflected on the child similarities by stating “*It seems like a real child because it is using trial and error with the words it uses, sees what happens and then express either happiness or disappointment. What is not similar is that it is not trying same word again, a real child would have kept using the same word until it got what it wanted.*” A final theme was **strategies** to approach the task. A few of the participants revealed trying to use some strategy to make the task easier and for learning, such as exclusion methods or “programming” the robot by intentionally giving certain objects for certain words. Finally, the simulated environment was generally negatively perceived as unnatural and game-like, making the task feel like a simulation; one participant said “*The design itself is very cute. It would be different with a real robot, a robot on a screen is more like a game rather than a real interaction.*”

Table 1. Mean Likert Scale Survey Results (1: Negative, 5: Positive)

Question	Mean (SD)
How did you experience the robot’s reactions when given an object it wanted?	4.67 (0.33)
How did you experience the robot’s visual response when given an object it wanted?	3.67 (1.03)
How did you experience the robot’s audible response when given an object it wanted?	4.67 (0.33)
How did you experience the robot’s response when given an object it didn’t want?	1.33 (0.52)
What was your impression of the robot’s appearance?	3.50 (1.29)

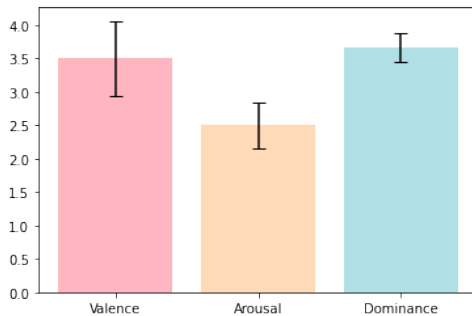


Fig. 4. Mean results (and standard error) of the Self Assessment Manikin Scale’s three dimensions of self-reported affective ratings. Each dimension is measured between 1 to 5 for valence (5=happy, 1=unhappy), arousal (5=stressed, 1=not stressed) and dominance (5=in control, 1=not in control)

5 Discussion and future work

This study investigated a human robot mutual learning setup where the robot’s ability to learn to associate words with internal needs, was determined by convergence rate and human experience was determined by self-reported affective ratings and a semi-structured interview and thematic analysis. The results, as shown in Figure 3, indicate that neither in the DOT nor the non-DOT condition did the robot achieve full convergence. It is, however, noteworthy that the DOT condition demonstrated a higher rate of learning compared to the non-DOT condition. This suggests that the robot’s learning process was influenced by the human’s learning being enhanced by the DOT, but further exploration with more iterations as well as more participants is necessary for conclusive results. Considering the low number of participants in this study, it would be of interest for further studies to assess whether the mean differences between DOT and non-DOT show a statistical significance in terms of both human correct responses (to the robot’s babblings) and robot language convergence rate. Investigating this DOT effect in the participants could provide further insights into how effective a setup like this could be for cognitive interventions, including memory training for people with dementia and MCI. Furthermore, the language acquisition model in this study is still in its infancy and thus quite simple and limited to a few internal needs. Future work will expand this model to better capture language learning and interaction complexities, making the model more generalizable and scalable, the current study is thus an important step in that direction.

The affective ratings, displayed in Figure 4, showed that participants reported high valence (happiness) and dominance (being in control), along with low arousal (not feeling stressed). The thematic analysis of the semi-structured interviews showed an emphasis on attitudes towards the setup, where some found the task relatively easy to comprehend, but others expressed a desire for a more meaningful and less simulated purpose behind the interaction. This suggests that incorporating a clearer objective or validation with the physical robot could enhance the user experience. Another theme was the negative and positive coding of the robot’s feedback and the results from the survey suggested that the positive audio feedback was more indicative of a positive response than the visual (movement) feedback. Fine-tuning the robot’s various signals could improve the user’s understanding of its feedback and thus also further enhance the setup, the DOT and its efficiency. Furthermore, the perception of the robot itself, including age, child-likeness and uncanniness were commonly occurring themes. The question of whether the robot being child-like facilitated the interaction or not remains to be investigated in future studies. A large amount of the socially assistive robots used for elderly care in particular tends to be designed to resemble either a pet or a child that needs to be cared for [27]. However, in the context of cognitive training and interventions the effectiveness of robot design in terms of age, anthropomorphism and social and affective role requires further investigation, e.g. through assessment of dimensions of warmth and competence [8]. Some participants attempted to strategize the robot’s learning process by using exclusion methods and intentionally providing certain objects for specific words. This finding demonstrates participants’ active engagement in the interaction and highlights the potential for further exploring different human-robot collaborative learning approaches/strategies.

In conclusion, the pilot study offers some promising results for the mutual learning human-robot interaction setup as proposed in this paper. The findings tentatively suggest that incorporating DOT could enhance both the human and the robot’s learning process. Moreover, participants generally had positive emotional experiences and expressed care and empathy towards the robot. To build upon these findings, future studies should consider increasing the number of iterations in the language acquisition process as well as the number of participants to evaluate the convergence rate and DOT-effect more comprehensively. Additionally, providing a clearer purpose or context to the interaction and refining the robot’s audiovisual feedback could lead to a more immersive and meaningful experience for users. Furthermore, investigating how users’ strategies and actions during the interaction impact the robot’s learning process could be of interest for collaborative human-robot learning systems. Finally, exploring interactions in more ecologically valid, i.e. physical environments settings with complex homeostatic drives, rather than a simulated environment and a robot with relatively simple homeostatic drives, may offer a more authentic and comprehensive understanding of human-robot interaction based on mutual learning with respect to homeostatic and affective robotic states. Ultimately, a target group for this work is persons with Mild Cognitive Impairment or dementia undergoing cognitive training based interventions. Given

further study and development of our approach, e.g. to include more complex drives and use of the physical robot, the mutual learning for a robot with homeostatic needs potentially provides: a) social (e.g. nurturing), b) ‘active’ (or co-) learning benefits to such participants that might provide positive therapeutic outcomes beyond more standard ‘passive’ learning practises.

References

1. Agüera-Ortiz, L., Pérez, R., Ruiz, F., JM, C.P., et al.: Social robots in advanced dementia. *Frontiers in Aging Neuroscience* **7**, 133–133 (2015). <https://doi.org/10.3389/fnagi.2015.00133>
2. Alnajjar, F., Khalid, S., Vogan, A.A., Shimoda, S., Nouchi, R., Kawashima, R.: Emerging cognitive intervention technologies to meet the needs of an aging population: a systematic review. *Frontiers in Aging Neuroscience* **11**, 291 (2019). <https://doi.org/10.3389/fnagi.2019.00291>
3. Alzheimer’s Society: Drug treatments and medication for alzheimer’s disease. <https://www.alzheimers.org.uk/about-dementia/treatments/dementia-drugs/drug-treatments-and-medication-alzheimers-disease> (2022), accessed: 2023-04-15
4. Andriella, A., Torras, C., Alenya, G.: Cognitive system framework for brain-training exercise based on human-robot interaction. *Cognitive Computation* **12**, 793–810 (2020). <https://doi.org/10.1007/s12559-019-09696-2>
5. Avila-Garcia, O., Cañamero, L.: Using hormonal feedback to modulate action selection in a competitive scenario. In: *From Animals to Animats 8: Proceedings of the 8th Intl. Conf. on Simulation of Adaptive Behaviour*, MIT Press, Cambridge, MA. pp. 243–52. Citeseer (2004). <https://doi.org/https://doi.org/10.7551/mitpress/3122.003.0031>
6. Bartoli, A., Catto, M., De Lorenzo, A., Medvet, E., Talamini, J.: Mechanisms of social learning in evolved artificial life. In: *Artificial Life Conference Proceedings 32*. pp. 190–198. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ... (2020)
7. Bradley, M.M., Lang, P.J.: Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* **25**(1), 49–59 (1994). [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
8. Carpinella, C.M., Wyman, A.B., Perez, M.A., Stroessner, S.J.: The robotic social attributes scale (rosas) development and validation. In: *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*. pp. 254–262 (2017). <https://doi.org/10.1145/2909824.3020208>
9. Cos, I., Canamero, L., Hayes, G.M., Gillies, A.: Hedonic value: Enhancing adaptation for motivated agents. *Adaptive Behavior* **21**(6), 465–483 (2013). <https://doi.org/10.1177/1059712313486817>
10. Davis, B.L., Macneilage, P.F.: Organization of babbling: A case study. *Language and speech* **37**(4), 341–355 (1994). <https://doi.org/10.1177/002383099403700401>
11. Esteban, L., Plaza, V., López-Crespo, G., Vivas, A.B., Estévez, A.F.: Differential outcomes training improves face recognition memory in children and in adults with down syndrome. *Research in Developmental Disabilities* **35**(6), 1384–1392 (2014). <https://doi.org/10.1016/j.ridd.2014.03.031>
12. Esteban, L., Vivas, A.B., Fuentes, L.J., Estévez, A.F.: Spatial working memory is enhanced in children by differential outcomes. *Scientific Reports* **5**(1), 17112 (2015). <https://doi.org/10.1038/srep17112>
13. eSwapp: Sharing is caring — why mutual learning is the workplace currency we all need to start spending (Mar 2021), <https://www.eswapp.com/mutual-learning/>
14. Fernandez-Lazaro, C.I., García-González, J.M., Adams, D.P., Fernandez-Lazaro, D., Mielgo-Ayuso, J., Caballero-Garcia, A., Moreno Racionero, F., Córdova, A., Miron-Canelo, J.A.: Adherence to treatment and related factors among patients with chronic conditions in primary care: a cross-sectional study. *BMC family practice* **20**(1), 1–12 (2019). <https://doi.org/10.1186/s12875-019-1019-3>
15. Hampel, H., Lista, S.: The rising global tide of cognitive impairment. *Nature Reviews Neurology* **12**(3), 131–132 (2016). <https://doi.org/10.1038/nrneurol.2015.250>
16. Jeong Hong, Y., Hye Jang, E., Hwang, J., Hoon Roh, J., Lee, J.H.: The efficacy of cognitive intervention programs for mild cognitive impairment: a systematic review. *Current Alzheimer Research* **12**(6), 527–542 (2015). <https://doi.org/10.2174/1567205012666150530201636>
17. Khan, I., Cañamero, L.: The long-term efficacy of “social buffering” in artificial social agents: Contextual affective perception matters. *Frontiers in Robotics and AI* **9**, 699573 (2022)
18. Kim, G.H., Jeon, S., Im, K., Kwon, H., Lee, B.H., Kim, G.Y., Jeong, H., Han, N.E., Seo, S.W., Cho, H., et al.: Structural brain changes after traditional and robot-assisted multi-domain cognitive training in community-dwelling healthy elderly. *PloS one* **10**(4), e0123251 (2015). <https://doi.org/10.1371/journal.pone.0123251>
19. Kohonen, T.: *Self-organization and associative memory*, vol. 8. Springer Science & Business Media (2012). <https://doi.org/https://doi.org/10.1137/1027085>

20. Lemhaouri, Z., Cohen, L., Cañamero, L.: The role of the caregiver’s responsiveness in affect-grounded language learning by a robot: Architecture and first experiments. In: 2022 IEEE International Conference on Development and Learning (ICDL). pp. 349–354. IEEE (2022). <https://doi.org/10.1109/ICDL53763.2022.9962197>
21. Lewis, M., Canamero, L.: Hedonic quality or reward? a study of basic pleasure in homeostasis and decision making of a motivated autonomous robot. *Adaptive Behavior* **24**(5), 267–291 (2016). <https://doi.org/10.1177/1059712316666331>
22. Lowe, D.G.: Object recognition from local scale-invariant features. In: Proceedings of the seventh IEEE international conference on computer vision. vol. 2, pp. 1150–1157. Ieee (1999). <https://doi.org/10.1109/ICCV.1999.790410>
23. Lowe, R., Almér, A., Gander, P., Balkenius, C.: Vicarious value learning and inference in human-human and human-robot interaction. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW). pp. 395–400. IEEE (2019). <https://doi.org/10.1109/ACIIW.2019.8925235>
24. Lowe, R., Almér, A., Lindblad, G., Gander, P., Michael, J., Vesper, C.: Minimalist social-affective value for use in joint action: a neural-computational hypothesis. *Frontiers in Computational Neuroscience* **10**, 88 (2016). <https://doi.org/doi.org/10.3389/fncom.2016.00088>
25. Lowe, R., Barakova, E., Billing, E., Broekens, J.: Grounding emotions in robots—an introduction to the special issue (2016). <https://doi.org/10.1177/1059712316668239>
26. Markelius, A., Sjöberg, S., Bergström, M., Ravandi, B.S., Vivas, A.B., Khan, I., Lowe, R.: Differential outcomes training of visuospatial memory: A gamified approach using a socially assistive robot. Submitted to *International Journal of Social Robotics* (2023)
27. Martinez-Martin, E., del Pobil, A.P.: Personal robot assistants for elderly care: an overview. *Personal assistants: Emerging computational technologies* pp. 77–91 (2018). https://doi.org/10.1007/978-3-319-62530-0_5
28. Mataric, M.J., Scassellati, B.: Socially assistive robotics. *Springer handbook of robotics* pp. 1973–1994 (2016). https://doi.org/10.1007/978-3-319-32552-1_73
29. McCormack, J.C., Elliffe, D., Virués-Ortega, J.: Quantifying the effects of the differential outcomes procedure in humans: A systematic review and a meta-analysis. *Journal of applied behavior analysis* **52**(3), 870–892 (2019). <https://doi.org/10.1002/jaba.578>
30. Peterson, G.B., Trapold, M.A.: Effects of altering outcome expectancies on pigeons’ delayed conditional discrimination performance. *Learning and Motivation* **11**(3), 267–288 (1980). [https://doi.org/10.1016/0023-9690\(80\)90001-6](https://doi.org/10.1016/0023-9690(80)90001-6)
31. Pieters, R., Racca, M., Veronese, A., Kyrki, V.: Human-aware interaction: A memory-inspired artificial cognitive architecture. *Cognitive Robot Architectures* **38** (2017)
32. Plaza, V., López-Crespo, G., Antúnez, C., Fuentes, L.J., Estévez, A.F.: Improving delayed face recognition in alzheimer’s disease by differential outcomes. *Neuropsychology* **26**(4), 483 (2012). <https://doi.org/10.1037/a0028485>
33. Rittmo, J., Karlsson, R., Gander, P., Lowe, R.: Minimalist social-affective value for use in joint action: a neural-computational hypothesis. *Acta Psychologica* **209**, 103134 (2020). <https://doi.org/doi.org/10.1016/j.actpsy.2020.103134>
34. Singh, A., Young, J.E.: A dog tail for utility robots: exploring affective properties of tail movement. In: *Human-Computer Interaction—INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part II* 14. pp. 403–419. Springer (2013). https://doi.org/10.1007/978-3-642-40480-1_27
35. Vivas, A.B., Ypsilanti, A., Ladas, A.I., Kounti, F., Tsolaki, M., Estévez, A.F.: Enhancement of visuospatial working memory by the differential outcomes procedure in mild cognitive impairment and alzheimer’s disease. *Frontiers in Aging Neuroscience* **10**, 364 (2018). <https://doi.org/10.3389/fnagi.2018.00364>
36. Walsh, C.A., Cahir, C., Tecklenborg, S., Byrne, C., Culbertson, M.A., Bennett, K.E.: The association between medication non-adherence and adverse health outcomes in ageing populations: a systematic review and meta-analysis. *British journal of clinical pharmacology* **85**(11), 2464–2478 (2019). <https://doi.org/10.1111/bcp.14075>
37. Yang, C.Y., Lu, M.J., Tseng, S.H., Fu, L.C.: A companion robot for daily care of elders based on homeostasis. In: 2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE). pp. 1401–1406. IEEE (2017). <https://doi.org/10.23919/SICE.2017.8105748>