Complex
Control
Systems

ISSN 1310-8255 Proceedings of the International Conference“Robotics & Mechatronics and Social Implementations”2018

TOWARDS SOCIAL COGNITIVE NEUROPSYCHOLOGY ACCOUNT OF HUMAN-ROBOT INTERACTION

Maya DIMITROVA[1*], Hiroaki WAGATSUMA[2], Vassilis KABURLASSOS[3], Aleksandar KRASTEV[4] , Ivan KOLEV[5]

[1)] Assoc. Prof., PhD, Interactive Robotics and Control Systems Department, IR-BAS, Sofia, Bulgaria
[2)] Assoc. Prof., PhD, Graduate School of Life Science and Systems Engineering, KYUTECH, Kitakyushu, Japan
[3)] Prof., PhD, Department of Computer & Informatics Engineering, EMaTTech, Kavala, Greece
[4)]Assoc. Prof., PhD, Interactive Robotics and Control Systems Department, IR-BAS, Sofia, Bulgaria
[5]MSc., PhD student, Sensors and Measurement Technologies in Robotics and Mechatronics, IR-BAS, Sofia, Bulgaria

Abstract: *The paper presents a novel framework towards the analysis of human-robot interaction, based on a new theoretical account called "social cognitive neuropsychology". This account is necessary to justify the introduction of complex technologies including humanoid and abstract robots in pedagogical settings, aiming also at special education. The proposed framework is supported by own published and non-published results from studies including humanoid robot NAO presenting a zoology lesson, supporting the socialising role of the robotic technology at school.*

Key words: *Humanoid robot NAO, education, pedagogical rehabilitation, intrinsic motivation, social motivation, socially aware robotics.*

1. INTRODUCTION

In recent years there is emerging evidence that robots (humanoid or abstract), as perceived by humans, can be considered a separate 'ontology of beings', capable of (semi-) autonomous performance and being easily re-programmed by their designers or users in order to adapt to, or influence, the current environmental and social situation in the World [1].

Many classifications of robotic technologies - as part of cyber-physical systems - exist, however these are mostly based on the mode of robot behaviour and functional capabilities - autonomous, mobile, degrees of freedom, appearance, operation tasks, etc. (e.g. [2], [3], [4]) rather than on the basic levels of influence on the human perception and emotionality, experienced by the used during the human-robot interaction process.

Current robotic technologies are endowed with cognitive abilities - sensing, perception, reasoning, decision-making, follow-up relearning and generating new knowledge based on previous experience (e.g. [5]). These are most often abstract and oriented towards solving logical tasks based on truthful assumptions, transferrable across domains, therefore domain-independent. The role of complex technologies as tools, assisting the process of socialisation of children has been extensively discussed, yet scarcely investigated in a systematic way. For example, teachers involved in the R-learning project in South Korea share the impression that robots can help children communicate more and in this way acquire better social skills[1].

A multi-agent system for teaching social skills to children with autism was designed and implemented in [6]. By agreeing on the task to allocate to a set of Kheppera robots, children playing in pairs learn to achieve better social skills, which are rewarded by robots performing as being told by the children. These and similar studies provide support to the pedagogical expectation about the wider application of robotic technology than just improving the problem solving or abstract thinking skills of children towards *understanding* the complexity of the surrounding World, which is largely socially mediated. The effect is being achieved in an implicit way, which relieves the burden on the child to self-control most of the time, which would provide a common ground for education of children with limited cognitive resource as well.

Robotic technology in schools is part of the cyber-physical system framework towards education. Unlike similar approaches, we derive the conceptualisation of the elements of the CPS from recent studies of brain processing underlying the social and cognitive behaviour in animals and humans. These studies serve as guidelines to translate outcomes from experimental research into directions for designing of improved and adapted to the human technologies and interfaces to those technologies.

An approach to incremental testing of the appropriateness of these novel technologies has also been proposed [7] and intended to be implemented further within the CybSPEED project.

[*] Corresponding author: IR-BAS, Akad. G. Bonchev Str., Bl. 2, Floor 4, Room 416, 1113 Sofia, Bulgaria, Tel.:+359 882 866 270

E-mail addresses: *maya.dimitrova.ir@gmail.com* (M.Dimitrova), *waga@brain.kyutech.ac.jp* (H. Wagatsuma), *vgkabs@teiemt.gr* (V. Kaburlasos), *aikrastev. iser.bas@gmail.com* (A.Krastev), *ivankolew91@gmail.com* (I. Kolev)

[1] r-learning.or.kr

2. CONCEPTUAL MODEL OF HUMAN-ROBOT INTERACTION

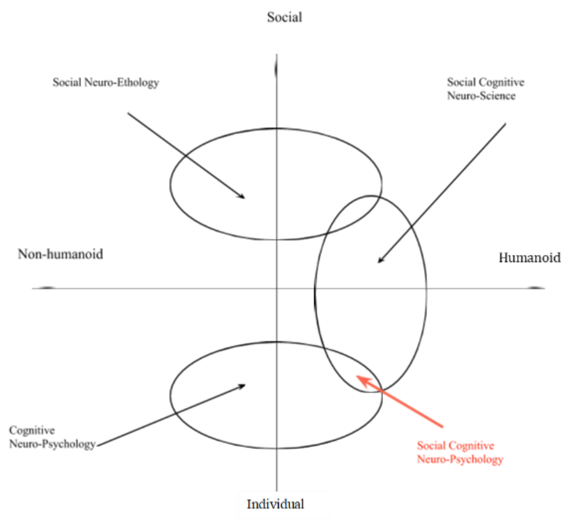

Fig. 1. Theories, potentially underlying the development of robotic solutions in education for helping people overcome learning difficulties.

2.1. Dimensions of human-robot interaction

The conceptual model of human-robot interaction in figure 1 deals with 2 proposed dimensions - physical resemblance to a human and socially-relevant reasoning abilities of robotic technologies and the related theoretical areas describing the underlying influences on perception and emotional reaction of humans to robots.

The social dimension spans cases, which define "social" literally - as an emergent relation between more than two participants - a small group, large group, very large group, population. Therefore, the opposite of "social" is not "non-social", but "individual" - any relation concerning one to one communication with another human or a robot.

Figure 1 presents 4 quadrants, starting from the bottom-left quadrant in clockwise direction. Quadrant I refers to cases of existing robotic technologies, characterised by abstract (non-humanoid) appearance and intended to support the individual at work, in the home, or in case of a special need. Theories of cognitive neuro-science provide relevant basis for modelling the underlying brain mechanisms of perception and emotional reaction to robotic technology, such as hemispheric specialisation [8], encoding and synchronisation of hippocampal and cortical activation, etc. [9].

Quadrant II refers to human-robot interaction cases inspired by the theory of social neuroethology (e.g. [10]). Social neuroethology investigates the neural mechanisms responsible for social behaviour in social animals such as ants, in most cases possessing features of physical non-resemblance to a human.

Quadrant III refers to the most investigated topic within the social robotics framework - implementation of humanoid robots participating in social situations with children, such as playing, communicating, singing and dancing together, taking professional roles on, etc. The relevant theories, underlying the modelling of human-robot interaction, are from the social-cognitive neuroscience, focused on revealing the neural mechanisms of empathy - mainly addressing the mirror neuron system and its relation to areas, responsible for the so called Theory of Mind (ToM) processing in humans and primates [11], [12].

Quadrant IV refers to the specific for the human brain processing inspiring the development of novel robotic technologies. It deals with examples of discriminative perception of emotional expressions irrespective of the nature of the face - human or machine [13], with distinctive features of perceiving a robot in comparison with the perception of a human [14], with the perception of robots with matching personalities to the human [15], etc. Based on such investigations we propose to delineate a special scientific area called "social-cognitive-neuropsychology".

Our theoretical hypothesis, motivating also the project CybSPEED, is that *the cognitive and emotional processing underlying human-robot interaction possesses all the potential for triggering the intrinsic motivations of children facing learning difficulties resulting in efficient compensation in learning* being the appropriate base of pedagogical rehabilitation via introducing robotic technologies in schools and day centres.

3. PHENOMENA IN HUMAN-ROBOT INTERACTION

3.1. Robot rejection by the human: Avoiding the Uncanny Valley phenomenon

Recent studies reveal that people perceive simultaneously multiple aspects of the agency attributed to the robot like, for example, "visceral factors of interaction", "social mechanics" and "social structures" [16] (p. 53). The authors relate the "visceral factors" to the 'uncanny valley' phenomenon, defined first by M. Mori (1970) – whenever the surface, physical attributes of the robot exceed a certain degree of resemblance to the human – feelings of unpleasantness and fear emerge in people communicating with the robots [17]. One of the hypotheses explaining it is that on a subtle discriminative, i.e. 'visceral', level, sensing of the 'strangeness' of robot behaviour emerges provoking negative reaction and thus presenting an obstacle to the flawless human-robot communication [18].

A novel hypothesis of the categorical nature of the 'uncanny valley' phenomenon has been proposed by R.K. Moore (2012) [19]. The advantage of his model is the mathematical description of the observed discrepancy between the subjective comfort of the interaction with the robot and the sudden repulsion by the realisation that the creature we are communicating with is a non-human. His model represents 'the human' as a normal distribution of 'objects' possessing features, defining the human as a conceptual category. This distribution is characterised by its mean, standard deviation and mathematical function that is descriptive of the form of the distribution and delimits the category boundaries. Representatives of the category 'human', sharing very typical or essential for the category features, in the abstract perceptual (i.e. internal) space are nearer the centre of the distribution, whereas the representatives with less typical (i.e. surface or characteristic) features are at the distribution outskirts – i.e. near the category boundaries. Inside the distribution, the probability of occurrence of a 'target' category – a human with typical 'human' features is higher, therefore there is better 'predictability' (as subjective anticipation) to encounter a typical human 'target'.

This distribution is combined by Moore with the distribution of the category 'non-human', which has bigger variance, hence broader span below the distribution function. The mean of the second - "background" - distribution does not typically coincide with the mean of the 'target' distribution in terms of the amount of features, describing essentially a 'human' in a categorical sense. However, if these coinciding features make the distribution means close enough, at some moment the forms of the overlapping functions, denoting category boundaries, form a 'function' with two optimums in Mori's sense – one positive when a humanoid robot resembles a stuffed animal - and one negative – when the interaction with the robot generates the feeling of communicating with a zombie.

The model of Moore plots 'affinity' in Mori's sense as familiarity plus 'perceptual tension'. The subtracting of the perceptual tension, playing the role of internal weighting factor, from the familiarity, will predict the observed phenomenon well. The model of the 'visceral uneasiness' of the human-robot interaction process, proposed by Moore, aims at explaining a variety of psychological phenomena when perceiving conflicting cues in an observed scene can invoke repulsion, anger or aggression and as such has deep societal validity.

An analysis has been performed in [20], for example, on the perception and emotional reaction during human-robot interaction on three conceptual levels - physical, social and psychological. Most of the existing intelligent systems attempt to predict behaviour in response to behaviour, i.e. to model the physical level of interaction between two interacting entities (humans and/or robots). A step higher in the reasoning abilities of the agent is its ability to predict behaviour in response to attitude, i.e. to model the social level of interaction between these interacting entities. Yet another possibility is the attempt to predict behaviour in response to opinion, i.e. to model the psychological level of interaction between these entities.

The psychological level of 'predicting behaviour in response to opinion', in our view, is the "uncanny" case. Whenever people react as if they feel that the behaviour of the robot is guided not just by attitude (social level), but by opinion (psychological level), by some kind of awareness like the one produced by a 'gaze sensor' [21], we expect to observe the 'uncanny valley' phenomenon. Robots need synthetic sensors like the 'gaze sensor' but they need not reinstate situations where the human 'gaze sensor' is on. They can rather reinstate feelings of positive attitude, friendliness, trust and compassion. Special questionnaires, distinguishing feelings close to perception of attitude from perception of opinion in human-robot interaction need to be designed to explore the validity of this hypothesis.

This focus of research, drawing a thin line between the social aspects of perceiving a humanoid or animated robot and the psychological effects that may be caused in the human-robot interaction process, is relevant to an overlapping section of theoretical research - the so-called "Social cognitive neuropsychology", which may provide guiding principles in formulating the educational aims of the future, based on the cyber-physical systems approach – more implicit that explicit instruction, entertainment, boosting intrinsic motivations in learning, joint individualised and socially-grounded approach to each child, reducing the stress and discouragement during learning and revealing the potential for cognitive and social development of the children/young adults, possibly vulnerable due to preexisting conditions, sharing a common environment with the neurotypical children and adults, as understood behind the phrase "pedagogical rehabilitation in education".

3.2. Socialisation vs. alienation: Truth and myths on the role of humanoid robots for socially mediated learning

We are exploring the scenario of a humanoid robot NAO playing the role of a teacher in several experimental trials. The aim is to investigate in parallel the social and cognitive motivation of human participants during a session of a zoology lesson, given by the robot (fig. 2). Cognitive motivation concept is similar to the recently proposed notion of "intrinsic motivation". The intrinsic motivation is the attraction of a cognitive system to novelty, making a difference between novelty and surprise.

Fig. 2. A set up for NAO teaching a zoology lesson.

Intrinsic motivations are being modelled in the FP7 project IM-CLeVeR [22], where agents – animals, humans and robots - are guided by internal drives for entertainment and socialisation being more sophisticated than the basic survival drives. IM-CLeVeR is embodying in i-Cub robot the intrinsic motivation of higher-level brains to seek new knowledge [23], thereby sustaining learning and self-improvement in the course of life.

Table 1

Level of recall of features of animals after a single trial

	Sea animals	**Forest animals**
M_{Total}	**1,59 (40%)**	**1,55 (39%)**
SD_{Total}	**0,89**	**1,03**

Our understanding of cognitive motivation emphasises the curiosity aspect of learning – when the presented material in the lesson is in itself rewarding for the student. Table 1 presents the amount of information learned after a single trial of a zoology lesson from [24].

The similar mean and SD_{Total} values of remembering information about the forest animals and the sea animals supports the conclusion that the set-up has provided a realistic reinstatement of the classroom situation where a humanoid robot NAO plays successfully the role of a teacher. In a follow up study we will investigate if the instruction to memorise the lesson will increase the amount of information learned from a single trial.

Table 2 presents the amount of time (%) devoted to viewing robots' hand than robot's face in the group of participants, wearing an AR eye-tracking device (11 participants).

Table 2

Amount of time (%) devoted to viewing robots' hand than robot's face in the AR group

	Robot's hand	Robot's face
M	42,90%	24,51%
SD	34,21%	31,49%

Students are attentive to the lesson and focus on the pointing gestures of the robot-tutor, which is in agreement with the expectation that pointing gestures and face movements are important to direct one's social attention, not just seeing it as a physical object. The amount of time viewing the robot hand did not differ significantly from the amount of time viewing the robot's face, as revealed by ANOVA, F(1, 20) =1, 56, p = 0,2255. In [25] it is shown that robots are not stereotypically defined by their face in a study comparing user attitude towards a human, a robot and a computer. This is in agreement with the present finding of feeling comfortable with a robot taking over a human profession like a teacher.

Social motivation in parallel with cognitive motivation was investigated by asking the participants in the study if they preferred the presence of a classmate during the lesson or were indifferent. Table 3 presents the obtained results, reported in [26].

Table 3

Percent positive comments and recommendations made by the socially motivated participants vs. the indifferent

	Positive comments	Recommendations
Socially-motivated	76,92%	61,54%
Indifferent	33,33%	77,78%

The participants who reported that they preferred the presence of a classmate during the session were 13 (the socially motivated group) in comparison with the indifferent ones, who were 9 (the socially indifferent group). The socially motivated group gave comparable amount of positive comments and recommendations as revealed by a single factor ANOVA, F(1, 24) = 0,69, p = 0,416. The socially indifferent group, however, gave significantly higher number of recommendations, than of positive comments as revealed by a single factor ANOVA, F (1,16) = 7,69, p = 0,014.

Therefore, we consider the expectation for the alienation role of the robots *a myth*, which seems to be contradicted by the experimental studies. Rather, our conceptualisation states, that people differ in their level of *social motivation* towards other people. Those, who tend to prefer other people's presence during the lesson, tend to prefer the robots, too. Whereas people, who are critical to other people, seem to be critical to the robots as well. This is an argument which is opposite of any claim that people, feeling distant from the others, may get attracted to the robots as substitute objects for human attachment.

It is expected to obtain further confirmation to the socialising role of robotic technology – either humanoid semi-humanoid or abstract, based on behavioural tests and brain research in relation to pedagogical rehabilitation in schools. Pedagogical rehabilitation is understood as a set of behavioural methods for teaching new learning and social skills, resembling games and classroom activities (rather than therapeutic approaches), encompassing the widest range of possible corrections of neurodevelopmental disorders like speech therapy, or focusing on minimal brain dysfunction, delays in acquisition of learning abilities, hyperactivity, attention deficit, etc. in an educational system's framework.

A promising novel trend of research on understanding the brain mechanisms of learning (in cognitive and social contexts) is "social cognitive neuroscience" (SCNS), currently providing evidence of the primary role of social interaction in the developmental process of shaping cognition (e.g. [27]). The social cognitive neuroscience forwards the idea of the emotion-cognition unity, where learning is driven by the rewarding role of the communication with the teacher and the peers.

We propose a further theoretical focus as the underlying conceptualisation of the human-robot interaction process in terms of "social cognitive neuropsychology" (SCNP). Whereas SCNS focusses on the understanding of the brain mechanisms, underlying social and socially mediated cognition, SCNP also provides insights on either the clinical relevance of robotic technology in social settings or of technology in the broader context of pedagogical and social communication in standard and special education [28].

It can also be modelled by the Lattice Computing account of (e.g. [29], [30]) for processing of both numerical and symbolic representation in different reinstatements of the human-robot interaction.

4. CONCLUSIONS

The paper presented a novel theoretical framework for implementation of complex socially-aware technologies like humanoid robots in standard and special education. Based on current social cognitive neuroscience research and on own studies is the proposal for a theory of "social cognitive neuropsychology" to be developed in order to support the novel approaches for introducing technology in education with emphasis on the technological enhancement of the process of pedagogical rehabilitation in special education. Novel experimental paradigms involving advanced brain computer and human computer interfaces are necessary to support the main ideas of the paper.

Acknowledgement. The authors acknowledge the financial support of H2020-MSCA-RISE-2017 for project No 777720 CybSPEED.

References

[1] M. Dimitrova, *Cognitive Theories for Socially-Competent Robotics in Education,* LAP Lambert Academic Publisher, 2016.

[2] D. Serpanos, *The Cyber-Physical Systems Revolution*, Computer, Vol. 51, No 3, 2018, pp. 70-73.

[3] V. Gunes, S. Peter, T. Givargis, F. Vahid, *A survey on concepts, applications, and challenges in cyber-physical systems*, KSII Transactions on Internet & Information Systems, Vol. 8, No. 12, 2014.

[4] Y. Zhang, M. Qiu, C. W. Tsai, M.M. Hassan, A. Alamri, *Health-CPS: Healthcare cyber-physical system assisted by cloud and big data*. IEEE Systems Journal, Vol. 11, No. 1, 2017, pp. 88-95.

[5] *The importance of iCub as a standard robotic research platform for embodied AI*, available at: `https://phys.org/news/2017-12-importance-icub-standard-robotic-platform.html` accessed: 2018-08-14.

[6] M. Dimitrova, N. Vegt, E. Barakova, *Designing a system of interactive robots for training collaborative skills to autistic children*, 15th IEEE International Conference on Interactive Collaborative Learning (ICL), 2012, pp. 1-8.

[7] M. Dimitrova, A. Lekova, S. Kostova, C. Roumenin, M. Cherneva, A. Krastev, I. Chavdarov, *A Multi-Domain Approach to Design of CPS in Special Education: Issues of Evaluation and Adaptation,* Proceedings of the 5th Workshop of the MPM4CPS COST Action, November 24-25, 2016, Malaga, Spain, 2016, pp. 196–205.

[8] C.E. Jung, L. Strother, D.J. Feil-Seifer, J.J. Hutsler, *Atypical Asymmetry for Processing Human and Robot Faces in Autism Revealed by fNIRS*, PLoS ONE 11(7): e0158804. `https://doi.org/10.1371/journal.pone.0158804`

[9] H. Wagatsuma and Y. Yamaguchi, *Cognitive Map Formation through Sequence Encoding by Theta Phase Precession*, Neural Computation, Vol. 16, No. 12, 2004, pp. 2665-2697.

[10] R. Boulay, S. Aron, X. Cerdá, C. Doums, P. Graham, A. Hefetz, T. Monnin, (2017). *Social life in arid environments: the case study of Cataglyphis ants*, Annual review of entomology, 62, 2017, pp. 305-321.

[11] L. Schilbach, *On the relationship of online and offline social cognition*, Frontiers in Human Neuroscience, Vol. 8, 2014, Article 278.

[12] L.T. Rameson, S.A. Morelli, and Matthew D. Lieberman, *The Neural Correlates of Empathy: Experience, Automaticity, and Prosocial Behavior*, Journal of Cognitive Neuroscience, Vol. 24, No.1, 2012, pp. 235-245.

[13] S. Dubal, A. Foucher, R. Jouvent, J. Nadel, *Human brain spots emotion in non humanoid robots.* Social, Cognitive and Affective Neuroscience, Vol. 6, No. 1, 2011, pp. 90-97.

[14] H. Admoni, C. Bank, J. Tan, M. Toneva, B. Scassellati, B. *Robot gaze does not reflexively cue human attention,* Proc. of the 33rd Annual Conf. of the Cognitive Science Society CogSci 2011, Austin TX USA, 2011, pp. 1983-1988.

[15] S. Andrist, B. Mutlu, A. Tapus, *Look Like Me: Matching Robot Personality via Gaze to Increase Motivation.* Proceedings of the 33rd annual ACM conference on human factors in computing systems, 2015, pp. 3603-3612.

[16] J.E. Young, J.Y. Sung, A. Voida, T. Igarashi, H.I. Christensen, R.E. Grinter, (2011) *Evaluating human-robot interaction. Focusing on the holistic interaction experience,* International Journal of Social Robotics Vol. 3, 2011, pp. 53-67.

[17] M. Mori, *Bukimi no Tani Genshō. The Uncanny Valley,* (K.F. MacDorman & T. Minato, Trans.) Energy, Vol. 7, 1970, pp. 33–35.

[18] F. MacDorman, *Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley.* Proc. ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science Vancouver Canada, 2006, pp.26–29.

[19] R.K. Moore, *A Bayesian explanation of the 'Uncanny Valley' effect and related psychological phenomena,* Nature Sci Rep 2, 864, doi:10.1038/srep00864, 2012.

[20] M. Dimitrova, H. Wagatsuma, *Designing humanoid robots with novel roles and social abilities*, Lovotics, Vol. 3, No 1, 2015.

[21] R.S. Jamisola, *Of Love and Affection and the Gaze Sensor*. Lovotics, Vol.1, No 1, 20104.

[22] *IM-CLeVeR* http://www.im-clever.eu/

[23] G. Baldassarre, T. Stafford, M. Mirolli, P. Redgrave, R. Ryan, A. Barto, *Intrinsic motivations and open-ended development in animals, humans, and robots: An overview.* Frontiers in Psychology, 5: 985, 2014 http://dx.doi.org/10.3389/fpsyg.2014.00985

[24] H. Wagatsuma, G.N. Tripathi, G. Ai, M. Dimitrova, A *Novel Experimental Framework for Testing Learner Attitude towards Humanoid Robot Tutoring Systems Based on Viewing Timeline Analysis* (unpublished).

[25] C.H. Ramey, C. H. (2006). *An inventory of reported characteristics for home computers, robots, and human beings: Applications for android science and the uncanny valley,* Proceedings of the ICCS/CogSci-2006 long symposium "Toward social mechanisms of android science", 2006, pp. 21-25.

[26] M. Dimitrova, H. Wagatsuma, G.N. Tripathi, G. Ai *Adaptive and Intuitive Interactions with Socially-Competent Pedagogical Assistant Robots*, Proc. 6th International Workshop on Interactive Environments and Emerging Technologies for eLearning (IEETeL 2015), 11-13 June, 2015, Lisbon, Portugal, 1-6.

[27] K.N. Ochsner, M.D. Lieberman, *The emergence of social cognitive neuroscience*, American Psychologist. Vol. 56, No. 9, 2001, pp. 17-34.

[28] L. Ozaeta, M. Graña, M. Dimitrova, A. Krastev, *Child oriented storytelling with NAO robot in hospital environment: preliminary application results*, Prob. Eng. Cybern. Robot. Vol. 69, pp. 21-29.

[29] V.G. Kaburlasos, C. Dardani, M. Dimitrova, A. Amanatiadis, M*ulti-robot engagement in special education: a preliminary study in autism*, IEEE International Conference on Consumer Electronics (ICCE), 2018, pp. 1-2.

[30] V.G. Kaburlasos, A. Kehagias, *Fuzzy inference system (FIS) extensions based on the lattice theory,* IEEE Transactions on Fuzzy Systems, Vol. 22, Issue 3, 2014, pp. 531-546.