Chapter 4
Learner Attitudes Towards Humanoid Robot Tutoring Systems:
Measuring of Cognitive and Social Motivation Influences

Maya Dimitrova
https://orcid.org/0000-0002-9975-1255
Bulgarian Academy of Sciences, Bulgaria

Gyanendra Nath Tripathi
https://orcid.org/0000-0001-9914-3789
Renesas Electronics, Japan

Hiroaki Wagatsuma
Kyushu Institute of Technology, Japan

Guangyi Ai
Neusoft Institute Guangdong, China

ABSTRACT

A novel framework for investigation of the learner attitude towards a humanoid robot tutoring system is proposed in the chapter. The theoretical approach attempts to understand both the cognitive motivation as well as the social motivation of the participants in a teaching session, held by a robotic tutor. For this aim, a questionnaire is delivered after the eye tracking experiment in order to record the type and amount of the learned material as well as the social motivation of the participants. The results of the experiments show significant effects of both cognitive and social motivation influences. It has been shown that cognitive motivation can be observed and analyzed on a very individual level. This is an important biometric feature and can be used to recognize individuals from patterns of viewing behaviors in a lesson. Guidelines, drawn from first-person accounts of learner participation in the study, are also formulated for achieving more intuitive interactions with humanoid robots intended to perform social jobs like being teachers or advisors.

DOI: 10.4018/978-1-5225-7879-6.ch004

INTRODUCTION

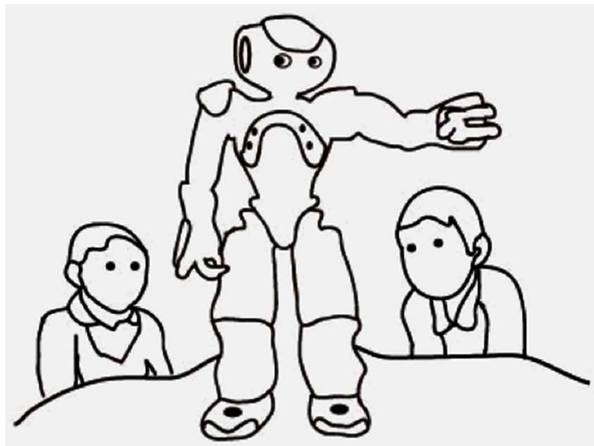
Trying to 'see' the robot tutor through the eyes of the learners by capturing the eye movements in an eye tracking experimental framework is a novel approach towards understanding the learners' attitude to inclusion of robots as teachers in the classroom. We propose a framework for testing the learner attitude towards a humanoid robot tutoring system that attempts to capture both the *cognitive motivation* as well as the *social motivation* of the participants in a teaching session, held by a robotic tutor. The important role of humanoid robots as technological support in the classroom has been extensively investigated recently both with typically developing children and children with learning difficulties (Lourens & Barakova; 2009; Krichmar & Wagatsuma, 2011; Leyzberg, Spaulding, Toneva, & Scassellati, 2012; Belpaeme et. al, 2013; Feng, Gutierrez, Zhang, & Mahoor, 2013; Huskens, Verschuur, Gillesen, Didden, & Barakova, 2013; Kim et. al, 2013; Anzalone et. al, 2014; Barakova, Kim, & Lourens, 2014). The aspects of being different from a human tutor and therefore more patient, less emotional, more amusing and drawing child's attention have been emphasized in these studies. The present study explicitly addresses the investigation of the 'social nature' of the learning process modulated by a humanoid robot in the classroom. Some results from the behavioral aspect of the current study were published in (Dimitrova, Wagatsuma, Tripathi, & Ai, 2015; Dimitrova, Wagatsuma, Kaburlasos, Krastev, & Kolev, 2018). Here we present in detail the novel experimental framework, proposed to investigate aspects of the attitude towards humanoid robot tutors combined with a 'viewing timeline analysis' (VTA) of eye-tracking data.

The inspiration for the experiment described in the paper has come from recent neuroscience research on the importance of the underlying brain mechanisms of 'social cognition' for shaping the cognitive abilities of the learner (Pfeiffer, Vogeley, & Schilbach, 2013; Schilbach, 2014). These studies can provide novel pedagogical insights towards designing robots to assist the teaching process in the classroom in order to achieve smooth and intuitive interaction of the robot with the learner at hand.

A promising novel trend of research on understanding the brain mechanisms of learning (in cognitive and social contexts) is "social cognitive neuroscience", which provides evidence of the primary role of social interaction in the developmental process of shaping cognition (Ochsner & Lieberman, 2001). M.D. Lieberman (2012) has proposed the concept of "social working memory" as distinct in its neurological basis from the commonly assumed cognitive 'working memory'. The areas of the brain that get involved while attending to a lesson, which is being explained with emphasis on its social relevance or historical context, are broader and involve the cognitive 'working memory' areas as well. Moreover, memories, created with the participation of 'social working memory areas' are much more durable than without them (which is actually a pedagogical aim by itself). Therefore, 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 or by the so called "intrinsic motivations" (e.g. Baldassarre & Mirolli, 2013), rather than by some functional self-realization notion. This requires designing innovative experimental paradigms to investigate the learners' attitudes towards humanoid robot tutoring systems, which aim at being more competent about the explicit and implicit aspects of the social communication process involved in education.

In respect to the issue of implementing social communication competence, K. Dautenhahn (2007) has summarized some of the existing definitions of a social robot and outlined five functional roles of being: "socially evocative", "socially situated", "sociable", "socially intelligent" and "socially interactive".

Figure 1. A humanoid robot tutoring framework implementing competence about the explicit and implicit aspects of the social communication process



According to Dautenhahn (2007, p. 584), the socially interactive robots can perform the following behaviors: "express and/or perceive emotions; communicate with high level dialogue; learn models of, or recognize other agents; establish and/or maintain social relationships; use natural cues (gaze, gestures, etc.); exhibit distinctive personality and character; and may learn and/or develop social competencies". Within this list of characteristic features of a 'social robot', the proposed experimental framework is focused on using 'natural cues' to 'develop social competencies' in the social situation of introducing a robot as a tutor in the classroom (Figure 1). The central point of the proposed novel experimental framework is measuring in *parallel* the *cognitive* and *social* motivation influences during learning from a humanoid robot tutor.

Cognitive Motivation Influences

Responses, triggered usually by physical stimulation, are considered cognitive in their nature and are processed by the mechanisms of overt attention under natural conditions (Taylor, Spehar, Donkelaar & Hagerhall, 2011; Kaspar & Konig, 2011; Borji, Sihite, & Itti, 2013). The human natural inclination to perceive, memorize and recollect physical events and objects from the surrounding environment comprises the spontaneous learning process. The curiosity, considered a fundamental orienting feature of human intelligence, is a feature that often can be attributed to a person (i.e. student) and has a gradation of being expressed in the classroom. Such aspects, guiding human behavior in the learning process, are manifestation of the *cognitive motivation* of the student to learn. Automatic processing such as *priming* or pre-exposure to a physical stimulus, event or objectified environment, belongs to this category of psychological processing, including perceiving people and their actions as physical objects (Oliva & Torralba, 2007; Sanada, Ikeda, Kimura, & Hasegawa, 2013; Dimiccoli, 2015). We refer to the analysis of the eye gaze behavior as to investigation of the *cognitive motivation* of the participants in the present experiment (Wykowska, Anderl, Schubo & Hommel, 2013).

Social Motivation Influences

In recent studies evidence is being accumulated about the neurological basis of social influences of the learning situation, within a novel theoretical framework called social-cognitive neuroscience in Ochsner & Lieberman (2001). A social motivation theory has been proposed by Chevallier, Kohls, Troiani, Brodkin, & Schultz (2012), stating that some aspects of the 'natural' human behavior are guided by seeking social contact and feeling pleasure solely by being in the company of other people. Novel studies reveal 'natural', i.e. 'hard-wired' in the neuronal activity processes such as "social perception" (e.g. Barraclough & Perrett, 2011), "social working memory" (Lieberman, 2012), "social gaze" of e.g. Emery (2000) and "social attention" of Langton, Watt, & Bruce (2000). It has been shown that the cues to the direction of social attention are not limited to eye gaze direction, but are combined with other cues, such as head orientation and pointing gestures. Moreover, it is concluded in Langton, Watt, & Bruce (2000, p. 56) that "…secondary cues, such as head orientation and pointing gestures, might provide more salient signal to the direction of another's attention than eye-gaze direction alone".

Viewing Timeline Analysis (VTL)

In the present study the fine grained eye gaze behaviors are investigated of the listeners to a lesson, delivered by a humanoid robot NAO, who is imitating the human teacher by using pointing gestures, head movement and eye-gaze direction when attempting an eye contact with the student. The point of interest is the relative amount of time spent on viewing these activities, performed by NAO, in order to propose a method called "Viewing timeline analysis" (VTL). The paper presents the main idea of this method, which will be developed in future studies.

The chapter presents a robotic-tutor scenario to investigate fine-grained gaze behavior of students when taught by a humanoid robot NAO, the obtained results from an experimental study and the design implications for pedagogical aims.

BACKGROUND

The experiments with unimodal and multimodal communication cues in human-robot interaction, presented by Torta, van Heumen, Piunti, Romeo, & Cuijpers (2015) are closely related to the present study. A NAO robot attempted to attract attention of a viewer of a TV show in a home environment. Three visual and one auditory actions were manipulated – NAO attempting to establish eye contact, waving gesture, blinking NAO's eyes and uttering the word "Hello". Reaction time was measured from the moment of noticing NAO in the periphery of one's vision by pressing the spacebar of a computer keyboard. Contrary to the expectations, reaction time was faster in the unimodal sound conditions than in the conditions of combined visual and sound effects to capture attention. In our study information was presented in a multimodal form – NAO explaining and pointing to an illustration. We expect to observe the normal student viewing behavior, on the one hand, and on the other – to try and explore comprehensively the fine grained gaze behavior of the student in the multimodal mode of information presentation. Hypothesis 1 is formulated in the next section based on the assumption that the dynamical behavior of the robot NAO is a holistic stimulus presenting a lesson in the classroom where multimodality is a precondition to the lesson presentation.

Eye tracking data is used to obtain deeper understanding of the actual visual processes during viewing NAO tutoring. In a study of Taylor, Spehar, Donkelaar & Hagerhall (2011) natural and artificial scenes as well as fractals and pink noise were presented successively. The results showed that the "biggest change occurred between initial and second observation and is expressed by a result pattern of increasing fixation durations, a decrease of saccade frequency and saccade length, and a reduced individual fixation distribution" (p. 4). We expect to obtain a similar pattern of saccadic eye movement change from the first to the second picture of the presented animals (based on their visual similarity) and formulated on this assumption Hypothesis 2.

In a study of Ai, Shoji, Wagatsuma, & Yasukawa (2014) participants viewed photographs with different figure-ground organization to extract profiles of spontaneous observation. Similarly, we want to extract profiles of gaze behaviors (as indicators of cognitive motivation) during viewing the actual robot and the illustrations of the lesson in the further analysis of the experimental data. Hypothesis 3 forwards the idea that the pattern of eye gaze can be assumed a biometric indicator which can bring new applications of the proposed approach. By applying the "viewing timeline analysis" (VTA) features can be extracted of individual viewing behavior for the purposes of the individualized tutoring in typical children and children with special needs.

Robotic reflexive cueing of attention was investigated by Admoni, Bank, Tan, Toneva, & Scassellati (2011) by implementing psychophysical methods. The participants viewed briefly on a laptop screen cues that were photographs of a human face, line drawing of a human face, arrow and two robotic faces with different degrees of anthropomorphism - Zeno (high) and Keepon (low), gazing in different directions. They were informed that a flashing stimulus will appear in the periphery of their viewing focus. The probability of the probe to appear in a location that is opposite to the cuing gaze was three times higher than to appear anywhere else on the screen, which was also explained to the participants. The theoretical hypothesis was that the response time will reflect the level of anthropomorphism of the cue. It was expected to observe that the information about probe appearing more often to the opposite of the gaze direction of the cue (human face, drawing of a human face, two types of robot faces or abstract drawing of an arrow) will result in different speed of responding to the probe, influenced by the level of human-likeness of the cue. This was not observed. Viewers' attention was reflexively cued by all stimuli except for the robotic faces. The robotic faces gaze did not interfere with the intentional orienting of the participants' gaze following the given instructions. Although this result may seem unexpected, it gives useful design clues for using robots in the classroom. We assume that the main potential of the robotic technology is in the ability to adapt the dynamics of its behavior to the most subtle human viewers' dynamics. The adaptive potential of the robot, especially in the time dynamics domain, is limitless and is of the kind that is not possible to expect from a human tutor. In the present study, we expect to observe robot cuing and capturing the attention of the viewer, which we attribute to the robot motion as temporal stimulus and use to formulate Hypothesis 4, unlike the static case of viewing snapshots of robot gaze in predicted directions as in (Admoni, Bank, Tan, Toneva, & Scassellati, 2011)

Not many studies attempt to relate cognitive and personality aspects of experiments on viewing behavior. In the study of Kaspar & Konig (2011) the participants in an experiment on focus of attention with an eye tracking system were delivered a personality test before the experiment. The test divided the participants into two groups of action- or state- oriented. A separate factor was introduced dividing the participants into a group, rating the viewed pictures as more interesting than the second group being less interested (by median splitting of the interestingness ratings). The main outcome was the three way interaction of the factors type of image viewed, action orientation and rating of the image. The action orientation in combination with high interestingness of the images had lowest saccade fixation and highest saccade frequency especially during viewing fractals or pink noise. The authors call it influence of "motivational disposition" on viewing behavior (p. 13). After rejection of pink noise images, the effects remained significant, which validated the subjective motivation influence on viewing behavior in general. However, it is important to note that the first two second's viewing revealed no effect of subject's interest in images on saccade length. Therefore, the subjective motivation is revealed at longer duration of viewing the images (6 sec) and may reflect the temporal dynamics of processing in the brain (Wagatsuma & Yamaguchi, 2001). Also, it seems that interestingness and action orientation may not be independent from each other, since they are revealed by similar questions. In our experiment, NAO took less than a minute to direct student attention to each of the pictures, so we did not expect differences due to the social motivation trait of the person. Rather, we expect independence relation (i.e. Dimitrova & Wagatsuma, 2011) between the cognitive and social motivation of a person, which can be considered features of the individual profile of the viewer and base the formulation of Hypothesis 5 on this assumption.

The aim of proposing a framework for relating the influences of the cognitive and the social motivation in studies on eye movement analysis is to extend the current approaches to extract features of scan path patterns (e.g. Kang & Landry, 2015) that are used to predict performance of novices vs. experts in aircraft conflict detection, for example. Another practical application is robot recognizing user focus of attention in order to support user intention by distinguishing sustained (longer duration) from shifted (brief) attention (e.g. Das, Rashed, Kobayashi, & Kuno, 2015). We propose an approach to predict the user's learning style, based on the analysis of the timeline of user viewing behavior for designing robot led learning courses tailored to the individual dynamics underlying the attentive processes (and to our knowledge is proposed for the first time). This approach can be extended even to special education, where robots can be pedagogical assistants to the teacher in the effort to overcome the learning limitations, which are due to scarce cognitive resource available to the individual learner.

AN EXPERIMENTAL DESIGN OF A NAO HUMANOID ROBOT TEACHER SCENARIO TO INVESTIGATE FINE-GRAINED STUDENT GAZE BEHAVIOR

The educational interaction process is assumed an interplay between joint attention and gaze-based interaction in a social communication scenario. A teacher captures the gaze of the child, then points to some illustrative material (joint attention), then gazes again at the child to find out if the point made has been understood by the learner. This constant interaction process is being paused periodically because the attention focus is tiring for the child and some rest is needed to accumulate some new attentional resource. The tutor, adapting to the dynamics of the child's attention, and the child, intuitively feeling comfortable with the tutor's teaching style, are both sides of a single educational process based on social communication competence at a very subtle level. This process is being reinstated in a NAO robot pedagogical framework in the present chapter.

Main Hypotheses of the Study

The following hypotheses are formulated within the proposed experimental design:

Hypothesis One: The fine grained gaze behavior of the student in the multimodal mode of information presentation will, in general, be consistent with the expectation that a robot tutor can be a successful substitute for a human tutor based on the analysis of eye tracking data, the recall test and the subjective reports.

Hypothesis Two: The pattern of saccadic eye movements change from the first to the second picture of the presented animals, based on their visual similarity, is expected to take place towards longer fixations as well as fewer and shorter saccades as a sign of transition from global to local processing.

Hypothesis Three: We expect to extract profiles of gaze behaviors (as indicators of cognitive motivation) during viewing the actual robot and the illustrations of the lesson in the analysis of the experimental data.

Hypothesis Four: We expect to observe robot cuing and capturing the attention of the viewer, which we attribute to the robot motion as a temporal stimulus, unlike the static case of viewing snapshots of robot gaze in predicted directions.

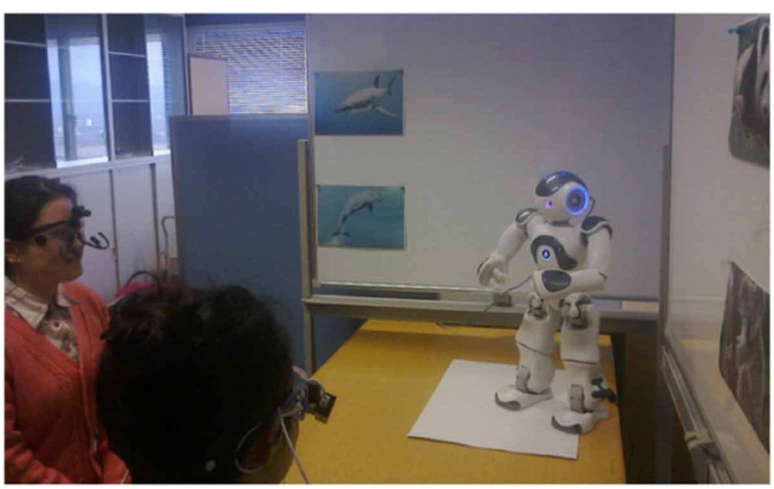
Hypothesis Five: We expect independence of the factors cognitive and social motivation, which can be considered features of the individual biometrics of the viewer. By establishing this, novel pedagogical strategies can be designed and implemented in a robot tutor for more adaptive and intuitive interaction with the current student.

Hypothesis Six: Cuing and capturing attention (indicators of learner attitude to the robot tutor) can be predicted based on a novel combination of features expressing the cognitive and social motivation for the individual learner by applying VTA.

Experimental Procedure

An experimental design was set up for demonstrating that a humanoid robot NAO can be used to modulate students' attention during a lesson. We investigated the spontaneous viewing behavior of typical students from University being taught by NAO by collecting eye tracking data (Figures 2 and 3).

Figure 2. Recording gaze behavior in a NAO tutor experimental framework using the AR system



Participants

Participants were eleven pairs of students and staff (17 male and 5 female students and staff) who knew English and were naïve to the purpose of the study. The average age was 32.6 years (23-55; SD = 9, 92).

The main selection criterion was to test viewing behavior of typical adults in order to obtain baseline levels of performance in human-robot context. Half of the participants had normal or corrected to normal vision via lenses (11 subjects). Having normal vision was not a prerequisite for the particular study because it tested a natural classroom situation with students adjusting their cognitive abilities to the multimodal style of the teacher. The other half of the participants wore glasses and took them off for the eye tracking part of the experiment. This did not change any of the dependent variables of the experiment.

Experimental Set Up

Figure 4 presents a diagram of the experimental setup with respect to the position, device and gender of the participants (blue for female). The yellow area represents the spatial organization of the experimental set up. The participants are invited in pairs and situated at 90 degrees to each other, and the robot is placed facing the diagonal line. In this way NAO can successively point out to one or the other board with displayed pictures of two classes of animals – forest animals on one of the boards and sea animals on the second.

Ethics

The study conformed to the national and EU guidelines for conducting experiments with human subjects. Written consent was obtained from all participants and they received a gift coupon as award after the experiment. The study was approved by the Ethics Committee of the Kyushu Institute of Technology, Japan, where the experiment took place.

Figure 3. Recording gaze behavior in a NAO tutor experimental framework using Takei Device

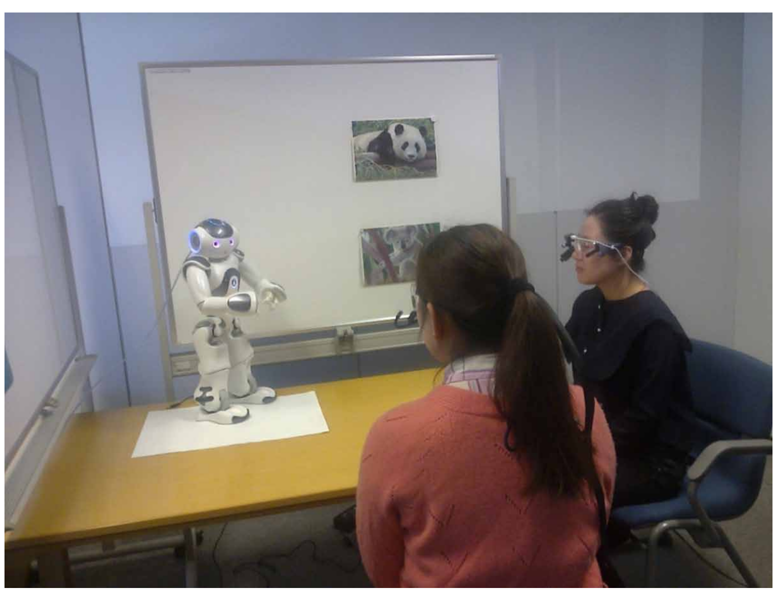
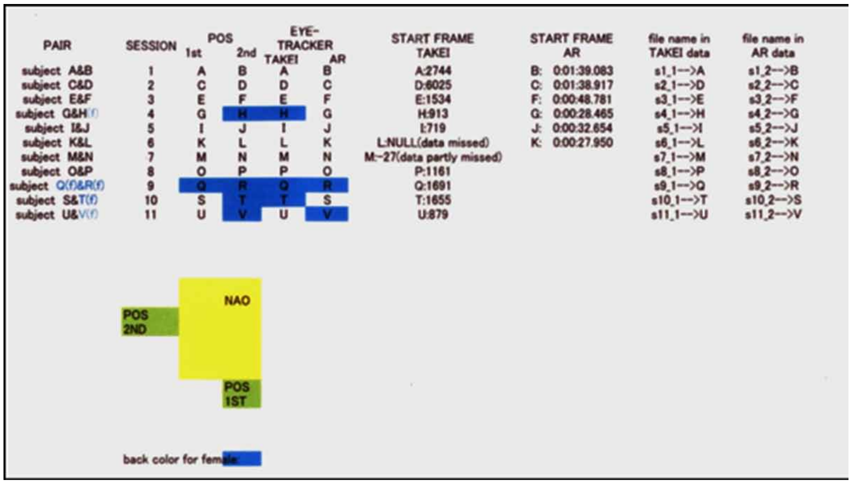


Figure 4. Diagram of the experimental setup with respect to position, device and gender of the participants (blue for female, white for male)



PAIR	SESSION	POS 1st	POS 2nd	EYE-TRACKER TAKEI	EYE-TRACKER AR	START FRAME TAKEI	START FRAME AR	file name in TAKEI data	file name in AR data
subject A&B	1	A	B	A	B	A:2744	B: 0:01:39.083	s1_1-->A	s1_2-->B
subject C&D	2	C	D	D	C	D:6025	C: 0:01:38.917	s2_1-->D	s2_2-->C
subject E&F	3	E	F	E	F	E:1534	F: 0:00:48.781	s3_1-->E	s3_2-->F
subject G&H(f)	4	G	H	H	G	H:913	G: 0:00:28.465	s4_1-->H	s4_2-->G
subject I&J	5	I	J	I	J	I:719	J: 0:00:32.654	s5_1-->I	s5_2-->J
subject K&L	6	K	L	L	K	L:NULL(data missed)	K: 0:00:27.950	s6_1-->L	s6_2-->K
subject M&N	7	M	N	M	N	M:-27(data partly missed)		s7_1-->M	s7_2-->N
subject O&P	8	O	P	P	O	P:1161		s8_1-->P	s8_2-->O
subject Q(f)&R(f)	9	Q	R	Q	R	Q:1691		s9_1-->Q	s9_2-->R
subject S&T(f)	10	S	T	T	S	T:1655		s10_1-->T	s10_2-->S
subject U&V(f)	11	U	V	U	V	U:879		s11_1-->U	s11_2-->V

Instructions and Behavioral Test

The instructions to the participants in the experiment are given in Table 1.

Apparatuses

For recording data of eye gaze behavior AR ViewPoint PC-60 Scene Camera with EyeFrame PCI digitizer was used, with scanning at 30 Hz. The system mainly provides positions of the gaze (x, y) and the height and width of the pupil of the right eye (Figure 2, a). As a second device was used Takei TalkEye Lite T.K.K. 2950 system with same frequency and recording x and y angles and pupil size (x, y) of the right eye (Figure 2, b).

Table 1. Instructions to the participants

No	Steps	Instructions
1.	Informed Consent Form	You are going to participate in a study on using the humanoid robot NAO as a tutor in the classroom. Imagine NAO is teaching a zoology lesson. We are interested in the spontaneous gaze behavior while watching a teacher. Please, make yourself comfortable in the chair. The session will take a few minutes. Please sign this form as consent to participate in the study.
2	Free Recall Test (random order of presentation)	1. Please, write down the main differences between a Shark and a Dolphin as explained by NAO. 2. Please, write down the main differences between a Panda and a Koala as explained by NAO.
3	Feedback on the Experiment	Please, write any comments if you have, about the NAO robot teacher
4	Social Motivation Question	Was it better to listen to NAO in the presence of another person or indifferent? (Please circle the correct answer) A) Better B) Indifferent

Figure 5 presents a snapshot of the visualization interface of the Takei TalkEye Lite T.K.K. 2950 system.

Procedure

The participants were given instructions to imagine NAO as a zoology teacher. NAO presented in a synthetic voice 7 sentences comparing two forest animals - panda and koala - and 7 sentences comparing two sea animals in English - shark and dolphin (Table 2). During pointing out to the first pair of animals, NAO was making eye contact with one of the participants, and during pointing out to the second – with the other participant.

The total time NAO presented each pair of animals was 40 seconds, so the learning session took 1 min 20 seconds in total with 10 seconds pause after each pair (memory consolidation time). The preparation phase with the eye tracking glasses adjustment took 5-6 min.

The text included features of animals of the following categories – geographical location, habits, appearance and species information (in this order). The photographs were taken from the respective entries

Figure 5. Visualization interface of the Takei TalkEye Lite T.K.K. 2950 system

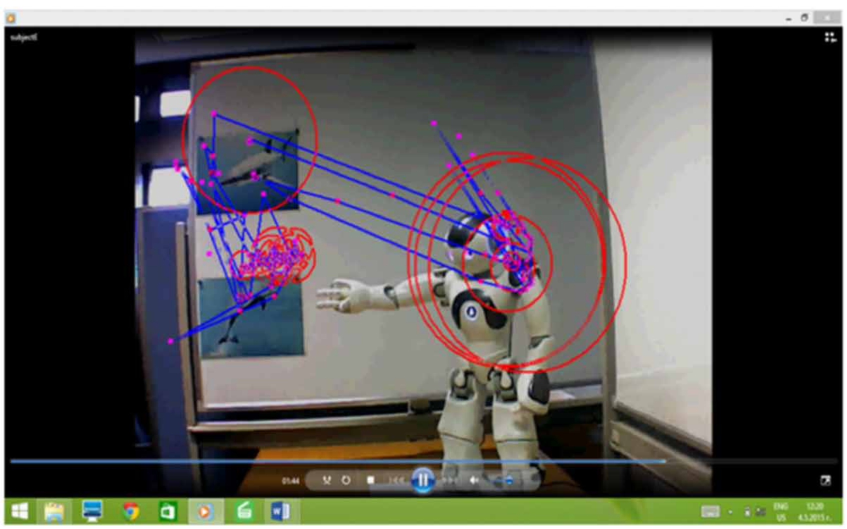


Table 2. Text pronounced by NAO

Condition	Text, Spoken by Nao about Forest Animals	Text, Spoken by Nao about Sea Animals
First animal presented and pointed at Text after making eye contact	This is a Panda. It leaves in China and eats bamboo leaves. Its color is white and black. Panda is a bear.	This is a Shark. It lives in the ocean and is dangerous. Its color is silver and white. Shark is a fish.
Second animal presented and pointed at Text after making eye contact	This is a Koala. It lives in Australia and eats eucalyptus leaves. Koala is not a bear.	This is a Dolphin. It lives in the sea and is not dangerous. Dolphin is not a fish.

in Wikipedia. The text about the appearance of the second presented animal was omitted in order to trace if previous knowledge would be used during recall. Due to the organization of the experiment, where pictures of the animals are presented like in a standard classroom – on a white board (instead on the computer screen) - it was not possible to exchange places and the order of presentation. Instead, the eye tracking systems were used in a counterbalanced order in respect to the presented animal photographs. The animal photographs were covered with a white sheet before the experiment.

After signing the informed consent form, the experimenter calibrated the eye tracking systems of the participants. During the calibration the participants were asked to focus at NAO's face, so using the robot as a vision focus became naturally. The photographs were uncovered and the experiment started. NAO stood up and pronounced the text about Panda while pointing at it and looking at it. Then it looked at the participant that was facing the forest animals and pronounced the rest of the text about Panda after making eye contact. Next NAO pointed out to the second forest animal (below the first one) and repeated the procedure. Then 10 seconds were allowed as memory consolidation time and NAO pronounced the text about the sea animals, addressing the participant, sitting facing the sea animals. In 10 more seconds' memory consolidation time, NAO pronounced the final words, waving its hand: "Have a good day! Bye!" and sat down. This was the end of the eye tracking part of the experiment.

Experimental Design

The experimental design is 2X2X2 factorial design with levels *type of attention modulation* (cuing vs. capturing attention*), type of communication signaling* (joint attention vs. eye-contact) and *stimulus exposure type* (exposed vs. pre-exposed stimulus). This design aims to capture the influences of the cognitive motivation to learn from NAO. The experiment investigates the *cognitive motivation and* the *social motivation* of the participants when taught by a robot in the presence of a classmate by asking a specially formulated social motivation question after the recall test (Question 4 in Table 2).

Independent and Dependent Variables

It was pre-tested if the type of attention modulation (cuing vs. capturing attention) interacted with the type of communication signaling (joint attention vs. eye-contact) in a 2X2 ANOVA. The type of attention modulation in terms of cuing vs capturing attention is related to the sitting position of the participant in respect to the position of the presented pictures and, in a way, represents the influence of a somewhat quasi-independent (or extrinsic) variable. Therefore, the independent variables of interest in the present experiment are the *type of communication signaling* (joint attention vs. eye-contact) and the *stimulus exposure type* (exposed vs. pre-exposed stimulus) and the dependent variables are the *level of recall* of features of the presented animals in the robot-performed zoology lesson (behavioral test) and *the nature of the saccadic eye movements* (cognitive test) – overall viewing time and individual viewing patterns.

Behavioral Test

After finishing the task, the participants were given a questionnaire, consisting of an unexpected free recall test, a question if they had comments about NAO as a robot teacher and a social motivation question if they were more comfortable in the presence of another person during a lesson delivered by a robot or indifferent.

RESULTS AND DISCUSSION

Results From the Behavioral Part of the Experiments

The behavioral part of the experiment was designed to provide a broader context to testing the operational hypotheses from One to Six. The results of the pre-test to see if the sitting position (view point) and type of stimulus (forest vs. sea animals) influenced the level of retention of the presented animal features in the zoology lesson are given in Table 3.

Table 3. Level of retention of the presented animal features in the zoology lesson

		Subject	Type of Stimulus (Forest vs. Sea Animals)			
			No	Sea Animals	No	Forest Animals
View point	Viewing sea animals first AR device T device	C G K O S		2 - habit, appearance 4 - appearance, habit, location, *species* 2 - habit, *species* 2 - habit 1 - species		1- location 4 - location, habit, appearance, *species* 3 location, appearance, *species* 2 location, appearance 1 species
		Mean_{AR} = Mean_{AR}/n = Mean_{AR} eye=		2,2 0,44 0,60		2,2 0,44 0,60
		A E I M Q U		1 – appearance 1 – *species* 1 – *species* 1 – *species* 1 – *species* 2 – location, habit		2 – location, habit 2 - appearance, habit 2 – location, habit - 1 – appearance -
		Mean_{T} = Mean_{T} /n= Mean_{T} eye =		1,4 0,23 0,67		1,4 0,23 0,00
	Viewing forest animals first AR device T device	B F J N R V		2 habit, *species* 1 *species* 4 appearance, habit, location, *species* 2 location, *species* 2 habit, *species* 1 habit		2 appearance, *species* 0 - 3 appearance, habit, location 2 location, *species* 2 appearance 0 -
		Mean_{AR} = Mean_{AR}/n = Mean_{AR} eye=		2,00 0,33 0,83		1,5 0,25 0,33
		D H L (f) P T		2 – habit, *species* 1 – location 1 – habit 1 – habit 1 – *species*		2 – location, habit 2 – appearance, habit 1 – location 2 – location, habit 1 - location
		Mean_{T} = Mean_{T} /n= Mean_{T} eye =		1,2 0,24 0,40		1,6 0,32 0,00
Mean_{Total} = SD_{Total} = Mean_{Total} /n=				1,64 0,90 0,41		1,59 1,05 0,40
M_{Total}eye = SD_{Total}eye =				0,64 0,49		0,23 0,43

The two-way ANOVA revealed a main effect of the type of sitting position used for viewing the lesson $F(1, 21) = 4{,}40$, $p = 0{,}0006$, but not type of the animals viewed first $(1, 21) = 0{,}06$, $p = 0{,}80$. The type of animal viewed first we have called influence of cued vs. captured attention. The first viewed animal is seen by captured attention by the robot pointing with its hand, and the second viewed animal is being cued by the robot pointing to the picture and looking at the other participant. Therefore it did not matter if the viewer was sitting frontal to the educational material that NAO was explaining, or was turning eyes to see what NAO was explaining. In this way a realistic classroom situation was reinstated with the robot-teacher being able to both cue and capture viewers' attention.

The remembered information after the eye contact was influenced by the attention being cued or captured in the total group of participants as revealed by a 2 way ANOVA, $F(1,21) = 14{,}54$, $p= 0{,}001$. The influence of the sitting position factor reached significance, too, $F(1,21) = 2{,}37$, $p = 0{,}027$. Since part of the participants on a sitting position wore one of the devices and part – the other, we tested the influence of the individual device on this pattern of results. The number of remembered animal features did not depend on the cueing vs. capturing attention in the group of subjects, sitting in position 1 (see figure 2) $F(1, 11) = 0{,}80$, $p = 0{,}39$, but depended on the type of device that was used for eye gaze recording AR vs. Takei $R(1,11) = 0{,}032$ (2 way ANOVA). Similar were the results from the 2 way ANOVA on the remembered items in the group of subjects sitting in position 2 with cueing vs. capturing attention being non-significant $F(1,8) = 1$, $p = 0, 35$, but type of device influencing the results, $F(1,8) = 8{,}5$, $p = 0{,}003$.

It seems that there was some problem with the Takei device for providing optimal view for the participants, so in our further analysis the examples are taken from the AR group. For example, the group wearing the Takei device and viewing the sea animals first recollected only the information from NAO making an eye contact, whereas the group wearing the Takei device and viewing the forest animals first did not recollect at all the information presented by NAO after making and eye contact. This can be due to the device itself being an obstacle to the proper vision of the participants. The pattern of data from the AR device favor the view that this is most probably an artefact from using the Takei device rather than a regularity in the group (see Table 4). In further experiments we recommend using the AR device as more reliable for the purpose of investigating the eye gaze behavior of students during a lesson, presented by a humanoid robot.

The pattern of recollection in the AR group was much more consistent than in the Takei group. Regarding the level of recollection of the information presented after the robot making an eye contact, the mean values of items remembered from the presented before and after making the eye contact were equal in the AR group, suggesting similar influences on the remembering of the information presented with robot pointing at pictures or during looking in the eyes of the listener. Therefore, the students felt comfortable as in a classroom where the teacher is using gaze behavior and pointing gestures to hold students' attention during the episode of presented information (which normally is made by the teacher between the natural pauses of the lesson).

As a primacy effect (see also Figure 6) the forest animals were less well remembered in the entire group. From the mean values of recall of sea and forest animals as well as the mean recalled after being presented after the eye contact of the robot with the students in the AR group it is evident that almost no influence is observed (i.e. values are similar in both conditions), which validates the experimental results in respect to the quasi-independent variable sitting position in the experimental 'classroom' being non-influential in the current experimental setting.

Table 4. Level of retention of the presented animal features in the zoology lesson in the AR group

<table>
<tr><th colspan="2" rowspan="2"></th><th rowspan="2">Subject</th><th colspan="4">Type of Stimulus (Forest vs. Sea Animals)</th></tr>
<tr><th>No</th><th>Sea Animals</th><th>No</th><th>Forest Animals</th></tr>
<tr><td rowspan="4">View point</td><td>Viewing sea animals first</td><td>C
G
K
O
S</td><td colspan="2">2 - habit, appearance
4 - appearance, habit, location, *species*
2 - habit, *species*
2 - habit
1 - species</td><td colspan="2">1- location
4 - location, habit, appearance, *species*
3 location, appearance, *species*
2 location, appearance
1 *species*</td></tr>
<tr><td>M =
M/n =
M eye =</td><td></td><td colspan="2">2,2
0,55
0,6</td><td colspan="2">2,2
0,55
0,6</td></tr>
<tr><td>Viewing forest animals first</td><td>B
F
J
N
R
V</td><td colspan="2">2 habit, *species*
1 *species*
4 appearance, habit, location, *species*
2 location, *species*
2 habit, *species*
1 habit</td><td colspan="2">2 appearance, *species*
0 -
3 appearance, habit, location
2 location, *species*
2 appearance
0 -</td></tr>
<tr><td>M=
M/n =
M eye =</td><td></td><td colspan="2">2,09
0,52
0,83</td><td colspan="2">1,56
0,39
0,33</td></tr>
</table>

Recollection of the Study Episode

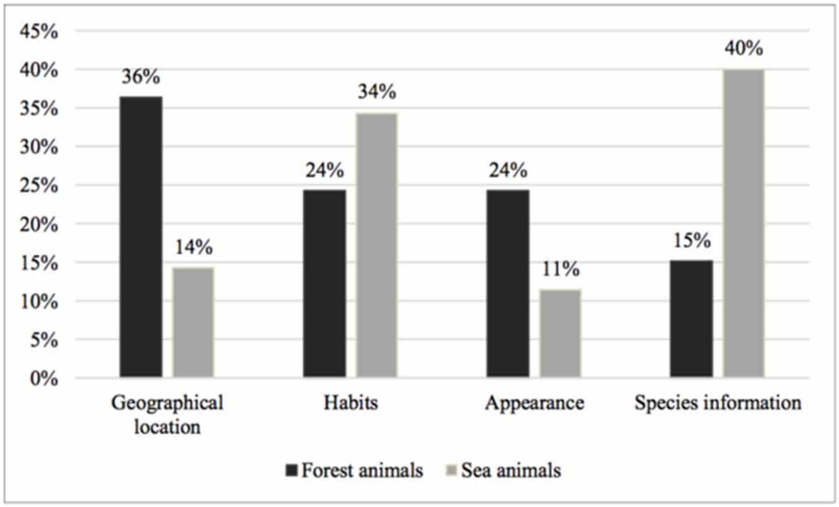
The participants were university students and staff and the teaching material was selected for children in early school. Table 5 presents the mean items recollected by the participants in total, being at the level of about 40% of the presented information.

The similar M_{Total} and SD_{Total} values of remembering information about the forest animals and the sea animals support the above conclusion that the setup has provided a realistic reinstatement of the classroom situation where a humanoid robot NAO plays successfully the role of a teacher. Figure 6 presents the proportions (in terms of percentage of the difference between animal features) recalled by the participants. It is important that they did not substitute the information, provided by NAO, with preexisting knowledge or from the visual memory. The standard primacy and recency effects are clearly observed: the geographical location of the forest animals was presented first, whereas the species information of the sea animals was presented last (the first and the last bars in the chart on recall in Figure 6). This validates the experimental results, as it is well known that information presented early during an experimental trial is well remembered as rehearsed longer than the rest, whereas the information, presented last stays in the working memory after the end of the trial (e.g. Baddeley, 1997).

Table 5. Mean feature items recollected by the participants in total

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

Figure 6. Recalled features per category (see text)



In the subjective reports the participants paid more attention to the human-robot interaction aspect of the study and commented how to improve robot performance to make it more intuitive. Some of the comments are related to the long 10 second pause after the presentations of the pairs of animals. The pause was made on purpose for memory consolidation after the presentation of the pairs of animals' text, so photos and mentioning it was omitted from the transcript of the protocols.

Social Motivation Results

Table 6 summarizes the comments of the participants regarding the robotic tutoring framework. Some participants gave positive comments and some commented on the performance of the actual robot as an imitation of a teacher. In overall, participants' comments addressed the way a robot interaction with the human should become more 'natural', i.e. more anthropomorphic as sound, posture, movement, etc. One explanation is that all participants were engineering and computer science students and staff, who were interested in designing user-friendly technological devices.

The observation of the participant behavior in the experiment demonstrates the relaxed way people perceive the option of having a robot teacher and the amusement this can bring to the educational process.

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). One particular result deserves special attention. Interestingly, the socially-motivated participants gave more positive comments and less recommendations to the robot teacher than the indifferent as illustrated in Figure7.

The type of device did not influence the number of positive comments or recommendations towards the humanoid robot teacher revealed by a 2 way ANOVA, $F(1,21) = 0{,}327$, $p = 0, 223$. 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 sig-

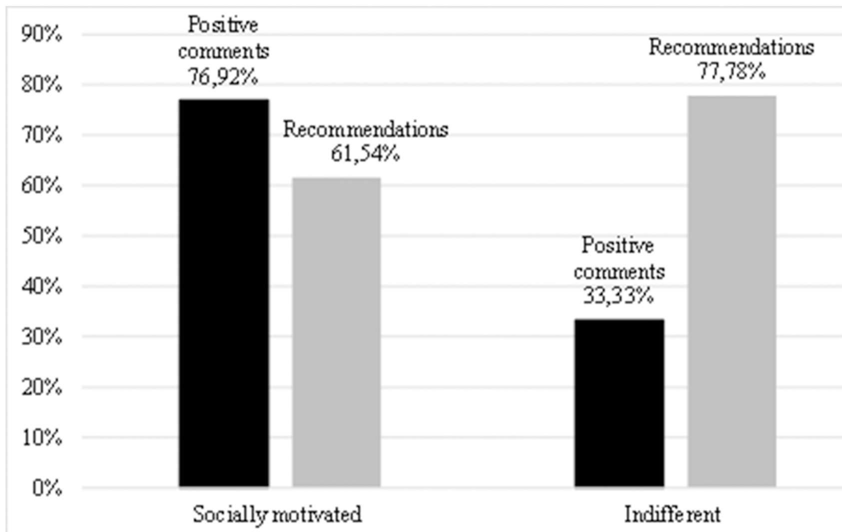
Table 6. Text pronounced by NAO

Parti-Cipant	Positive Comment	Recommendation
A		"May lip movements and more human-like gestures can help in the retention".
B		"The direction to the picture it points is not accurate. Sometimes if user has no knowledge about pictures may find hard to distinguish which one it is pointing to."
D	"Interesting to listen from a robot teacher."	
E	"I like the flashing light on the robot."	"May need some more movement for more interesting attraction."
F	"Very good of teaching."	"Sometimes voice in not much understandable."
G	"It is very nice to look at. When it makes eye contact, its influence is huge."	"Maybe its going too fast."
H	"It is easy for me to understand."	
I	"The NAO robot is a good teacher."	
J	"Interested to know how it follows the motion of the human."	
K	"The explanations were nice," but the interval between them was unnatural.	
L	"It was awesome to experience listening NAO robot teacher"	
M		"I think speaking (can be) more clearly."
N		"I heard some audio noise on his/her voice, which was a little bit disturbing to listen to his/her voice."
O	"NAO robot is an intelligent robot."	
P	"Nearly very good."	"If it is possible more clear speaking form NAO."
Q	"Robot teacher is very good, but …"	"… speaking is not clear."
R	"It is more interesting to listen to NAO" "And if it can say some jokes like person, I think it will make the phenomenon more comfortable and relaxible."	"I think it can explain more and more active" Sometimes it can point to the right direction of the picture"
U		"Well, it's better to give instruction by robot to listen carefully."
V		"NAO's way of speaking was a bit unnatural."

nificantly higher number of recommendations, than of positive comments as revealed by a single factor ANOVA, F(1,16) = 7,69, p = 0,14.

The finding that *the socially-motivated participants were more positive overall* towards the robot tutor than those who were socially-indifferent deserves further investigation with more participants, yet it provides support to the proposed in the present chapter framework that social interaction is an essential (and broader) concept for education than making mere attempts for memorizing the contents of the lesson. It confirms in a behavioral setting the 'social working memory' proposal of M.D. Lieberman (2012). With a human teacher, people are usually more attentive to the actual information presented in verbal and pictorial forms. With NAO tutoring it seems that the participants focus on the entertaining and the communicative aspects of the human-robot interaction process, rather than on the information

Figure 7. Percent positive comments (dark bar) and percent recommendations (light bar) made by the socially motivated participants vs. the indifferent



that NAO is giving by speaking and by pointing to pictures. This is also found in the subjective reports after the session with NAO.

Analysis of the Eye Tracking Data

Figure 9 presents example of a pie chart obtained about time devoted to viewing the robot or the pictures, demonstrated by the robots. Figure 10 present an example of the timeline diagrams of the viewing behavior. Both represent the data received from the AR ViewPoint PC-60 Scene Camera. It was evident that there is no dependency between any indicators of the social and cognitive motivation. Therefore, these two aspects of user behavior towards a robot appear orthogonal and fit well with a novel human-robot interaction account, where the *cognitive* and *social* motivation of a person can be considered independent dimensions. For a human-robot system, diagnostic of the user, this finding is important in order to make possible to draw a relevant profile of each individual user and be able to predict behavior based on both qualitative and quantitative indicators.

Testing the Individual Hypotheses

Hypothesis One: The fine grained gaze behavior of the student in the multimodal mode of information presentation will, in general, be consistent with the expectation that a robot tutor can be a successful substitute of a human tutor based on the analysis of eye tracking data, the recall test and the subjective reports. Table 7 displays the amount of time devoted to viewing robots' hand and robot's face in the AR group.

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

Table 7. Amount of time (%) devoted to viewing robots' hand vs. robot's face in the AR group

	Robot's Hand	Robot's Face
M_{AR}	42,90%	24,51%
SD_{AR}	34,21%	31,49%

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, $F(1,20) = 1,56$, $p = 0,2255$. 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 (Ramey, 2006). This is in agreement with the present finding of feeling comfortable with a robot taking over a human profession like a teacher.

Hypothesis Two: The pattern of saccadic eye movements change from the first to the second picture of the presented animals, based on their visual similarity, is expected to take place towards longer fixations as well as fewer and shorter saccades as a sign of transition from global to local processing. Figure 8 presents the amount of time devoted to viewing the first presented animal compared to the second in the AR group.

The time spent viewing 4 pictures of animals differed from one to another in the AR group as revealed by one way ANOVA, $F(3, 40) = 6,29$, $p = 0,001$. The time spent viewing Panda differed significantly from the time spent viewing Koala as revealed by ANOVA, $F(1, 20) = 8,64$, $p = 0,008$. The time spent viewing Shark did not differ significantly from the time spent viewing Dolphin as revealed by ANOVA $F(1, 20) = 0, 36$, $p = 0,5556$ (see Figure 8).

Figure 8. Pre-exposure effect of a similar category

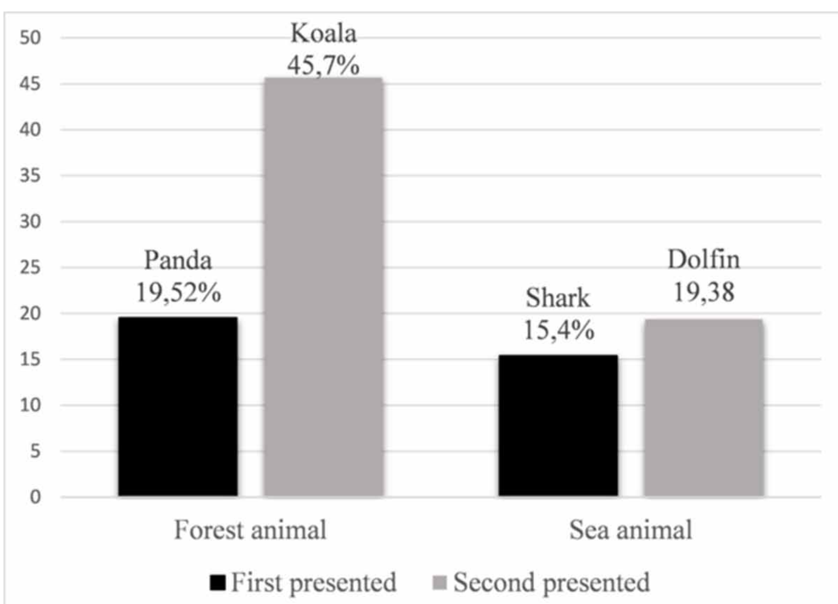


Figure 8 illustrates the pre-exposure effect of the similar category for both types of animals. As expected, the pre-exposure to Panda leads to bigger time devoted to viewing Koala and the pre-exposure to Shark leads to bigger time devoted to viewing Dolphin (although insignificant in the second case).

The pie chart in Figure 9 represents example of the individual proportions of viewing times in the AR group. Subjects B, C and J (Figure 4) are the group devoting most of their attention to viewing the pictures, whereas the rest devote most of their attention to viewing the robot. No other similarities can be observed among them. Evidently, cognitive motivation can be observed and analyzed on a very individual level in a way similar to reading the fingerprints of a person. This is an important biometric feature and can be used to recognize individuals based on patterns of viewing behaviors in a lesson, as well as to predict motivation to focus on the lesson via VTA ("viewing timeline analysis").

Robot cuing and capturing attention is seen from the pattern of individual timelines of viewing behavior (Figure 10). Subjects followed the pointing behavior of the robot from one picture to another and spent almost all of the time in viewing the teacher and the lesson. Formal analyses will be performed in future studies. The final sentence of the presented information about each animal is made in a mode of robot attempting eye contact. With a robot teacher, students behaved in full awareness that the robot is not a substitute for a human, but is performing a human role of a professional for a special reason. This is a conclusion that has to be accounted for when designing humanoid robots for playing different roles in human society that are not attempting to substitute the humans, but to assist them in their professional work.

Figure 9. Example pie chart of % individual viewing time during the robot lesson

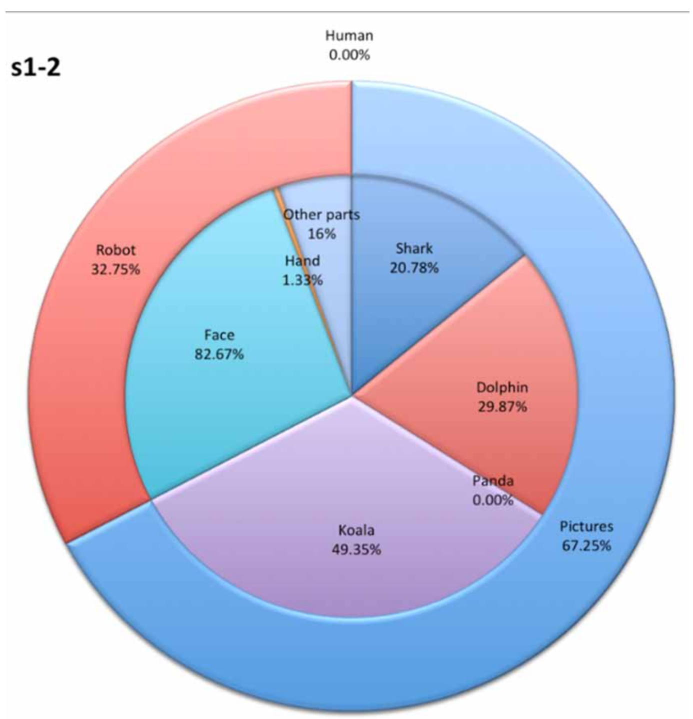
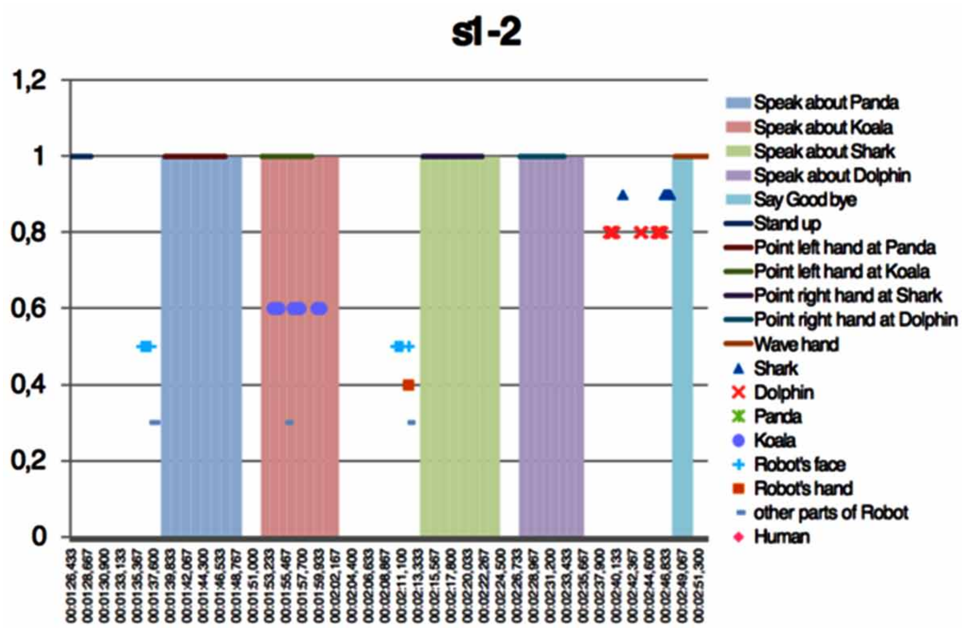


Figure 10. Example of a personal timeline of a learner viewing Panda first



CONCLUSION AND FUTURE WORK

With the inclusion of robots in the educational and instructional activities of all possible aspects of life in the near future, and with the increased cognitive competence implemented in technology, people of different learning styles and educational needs will rely on learner-adaptive instruction and intuitive human-robot interface. This is particularly valuable if the robotic technology is implemented in adaptive and intuitive learning settings for children with special educational needs, but also in the classroom in general.

The eye tracking data revealed the expected effects of attention cueing of the first participant to the first pair of animals, and of the second participant to the second pair of animals. Also the expected effect of attention capturing of the second participant by pointing to the first pair of animals, or attention capturing of the first participant to the second pair of animals was observed.

As a novel finding we established the participation of *both* social motivation and cognitive motivation in the process of perceiving a robot tutor. People demonstrated main concern with robot social communication competence/ability before actually paying attention to the contents of the lesson to be learned. This is an outcome that has to be taken into account when designing novel robotic tutoring systems.

It has become evident in this study that cognitive motivation can be observed and analyzed on a very individual level in a way similar to reading the fingerprints of a person. This is an important biometric feature and can be used to recognize individuals from patterns of viewing behaviors in a lesson for robotic implementation. Future work will include tests of the VTA approach and implementation of the proposed framework in different educational settings.

ACKNOWLEDGMENT

This work is partially supported by Grant 777720 of H2020-MSCA-RISE-2017 for project CybSPEED (2017 - 2021), MEXT/JSPS KAKENHI 15H05878, 16H01616, 17H06383 and the New Energy and Industrial Technology Development Organization (NEDO). The actual experiment was performed in 2015 while Maya Dimitrova was on a FY2014 JSPS Invitation Fellowship Program (Short-Term) for Research in Japan ID No. S-14156 at Kyushu Institute of Technology (KYUTECH). The authors are grateful to MSc Mayu Ichiki, who encoded the data for the timeline analysis.

REFERENCES

Admoni, H., Bank, C., Tan, J., Toneva, M., & Scassellati, B. (2011, January). Robot gaze does not reflexively cue human attention. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (*Vol. 33*, No. 33). Academic Press.

Ai, G., Shoji, K., Wagatsuma, H., & Yasukawa, M. (2014). A Structure of Recognition for Natural and Artificial Scenes: Effect of Horticultural Therapy Focusing on Figure-Ground Organization. In *Advanced Intelligent Systems* (pp. 189–196). Cham: Springer. doi:10.1007/978-3-319-05500-8_18

Anzalone, S. M., Tilmont, E., Boucenna, S., Xavier, J., Jouen, A. L., Bodeau, N., ... Chetouani, M. (2014). How children with autism spectrum disorder behave and explore the 4-dimensional (spatial 3D+ time) environment during a joint attention induction task with a robot. *Research in Autism Spectrum Disorders*, *8*(7), 814–826. doi:10.1016/j.rasd.2014.03.002

Baddeley, A. D. (1997). *Human Memory: Theory and Practice*. Psychology Press.

Baldassarre, G., & Mirolli, M. (2013). Intrinsically motivated learning systems: an overview. In *Intrinsically Motivated Learning in Natural and Artificial Systems* (pp. 1–14). Berlin: Springer. doi:10.1007/978-3-642-32375-1_1

Barakova, E. I., Kim, M. G., & Lourens, T. (2014, June). Development of a robot-based environment for training children with autism. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 601-612). Springer. 10.1007/978-3-319-07446-7_58

Barraclough, N. E., & Perrett, D. I. (2011). From single cells to social perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1571), 1739–1752. doi:10.1098/rstb.2010.0352 PMID:21536557

Beach, P., & McConnel, J. (2018). Eye tracking methodology for studying teacher learning: A review of the research. *International Journal of Research & Method in Education*, 1–17. doi:10.1080/1743727X.2018.1496415

Belpaeme, T., Baxter, P., De Greeff, J., Kennedy, J., Read, R., Looije, R., ... Zelati, M. C. (2013, October). Child-robot interaction: Perspectives and challenges. In *International Conference on Social Robotics* (pp. 452-459). Springer. 10.1007/978-3-319-02675-6_45

Borji, A., Sihite, D. N., & Itti, L. (2013). What stands out in a scene? A study of human explicit saliency judgment. *Vision Research*, *91*, 62–77. doi:10.1016/j.visres.2013.07.016 PMID:23954536

Cantoni, V., Galdi, C., Nappi, M., Porta, M., & Riccio, D. (2015). GANT: Gaze analysis technique for human identification. *Pattern Recognition*, *48*(4), 1027–1038. doi:10.1016/j.patcog.2014.02.017

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239. doi:10.1016/j.tics.2012.02.007 PMID:22425667

Das, D., Rashed, M. G., Kobayashi, Y., & Kuno, Y. (2015). Supporting human–robot interaction based on the level of visual focus of attention. *IEEE Transactions on Human-Machine Systems*, *45*(6), 664–675. doi:10.1109/THMS.2015.2445856

Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1480), 679–704. doi:10.1098/rstb.2006.2004 PMID:17301026

Diehl, J. J., Crowell, C. R., Villano, M., Wier, K., Tang, K., & Riek, L. D. (2014). Clinical applications of robots in autism spectrum disorder diagnosis and treatment. In *Comprehensive Guide to Autism* (pp. 411–422). New York, NY: Springer. doi:10.1007/978-1-4614-4788-7_14

Dimiccoli, M. (2016). Figure–ground segregation: A fully nonlocal approach. *Vision Research*, *126*, 308–317. doi:10.1016/j.visres.2015.03.007 PMID:25824454

Dimitrova, M., & Wagatsuma, H. (2011). Web agent design based on computational memory and brain research. *Information Extraction from the Internet*, 35-56.

Dimitrova, M., Wagatsuma, H., Kaburlasos, V., Krastev, A., & Kolev, I. (2018). Towards Social Cognitive Neuropsychology Account of Human-Robot Interaction. *Complex Control Systems*, *1*, 12-16. Retrieved from http://ir.bas.bg/ccs/2018/2_dimitrova.pdf

Dimitrova, M., Wagatsuma, H., Tripathi, G. N., & Ai, G. (2015, June). Adaptive and intuitive interactions with socially-competent pedagogical assistant robots. In *Proc. 6th International Workshop on Interactive Environments and Emerging Technologies for eLearning (IEETeL 2015)* (pp. 1-6). IEEE. 10.1109/ITHET.2015.7218031

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, *24*(6), 581–604. doi:10.1016/S0149-7634(00)00025-7 PMID:10940436

Feng, H., Gutierrez, A., Zhang, J., & Mahoor, M. H. (2013, September). Can NAO robot improve eye-gaze attention of children with high functioning autism? In *2013 IEEE International Conference on Healthcare Informatics* (pp. 484-484). IEEE. 10.1109/ICHI.2013.72

Huskens, B., Verschuur, R., Gillesen, J., Didden, R., & Barakova, E. (2013). Promoting question-asking in school-aged children with autism spectrum disorders: Effectiveness of a robot intervention compared to a human-trainer intervention. *Developmental Neurorehabilitation*, *16*(5), 345–356. doi:10.3109/17518423.2012.739212 PMID:23586852

Ichiki, M., Ai, G., Ooi, J. S., & Wagatsuma, H. (2016). A Comparative Analysis of Indexing of Mental Workload by using Neuro-Driving Tools based on EEG Measurements Coupling with the Eye-Tracking System. *Frontiers in Neuroinformatics*. *Conference Abstract: Neuroinformatics*, *2016*. doi:10.3389/conf.fninf.2016.20.00041

Kang, Z., & Landry, S. J. (2015). An eye movement analysis algorithm for a multielement target tracking task: Maximum transition-based agglomerative hierarchical clustering. *IEEE Transactions on Human-Machine Systems*, *45*(1), 13–24. doi:10.1109/THMS.2014.2363121

Kaspar, K., & König, P. (2011). Overt attention and context factors: The impact of repeated presentations, image type, and individual motivation. *PLoS One*, *6*(7), e21719. doi:10.1371/journal.pone.0021719 PMID:21750726

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1038–1049. doi:10.100710803-012-1645-2 PMID:23111617

Krichmar, J. L., & Wagatsuma, H. (Eds.). (2011). *Neuromorphic and Brain-Based Robots*. Edinburgh, UK: Cambridge University Press. doi:10.1017/CBO9780511994838

Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*(2), 50–59. doi:10.1016/S1364-6613(99)01436-9 PMID:10652522

Leyzberg, D., Spaulding, S., Toneva, M., & Scassellati, B. (2012, January). The physical presence of a robot tutor increases cognitive learning gains. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (*Vol. 34*, No. 34). Academic Press.

Lieberman, M. D. (2012). Education and the social brain. *Trends in Neuroscience and Education*, *1*(1), 3–9. doi:10.1016/j.tine.2012.07.003

Lourens, T., & Barakova, E. (2009, June). My sparring partner is a humanoid robot. In *International Work-Conference on the Interplay between Natural and Artificial Computation* (pp. 344–352). Berlin: Springer. doi:10.1007/978-3-642-02267-8_37

Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *The American Psychologist*, *56*(9), 717–734. doi:10.1037/0003-066X.56.9.717 PMID:11558357

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*(12), 520–527. doi:10.1016/j.tics.2007.09.009 PMID:18024143

Pfeiffer, U. J., Vogeley, K., & Schilbach, L. (2013). From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience and Biobehavioral Reviews*, *37*(10), 2516–2528. doi:10.1016/j.neubiorev.2013.07.017 PMID:23928088

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

Sanada, M., Ikeda, K., Kimura, K., & Hasegawa, T. (2013). Motivation enhances visual working memory capacity through the modulation of central cognitive processes. *Psychophysiology*, *50*(9), 864–871. doi:10.1111/psyp.12077 PMID:23834356

Schilbach, L. (2014). On the relationship of online and offline social cognition. *Frontiers in Human Neuroscience*, *8*, 278. doi:10.3389/fnhum.2014.00278 PMID:24834045

Taylor, R., Spehar, B., Hagerhall, C., & Van Donkelaar, P. (2011). Perceptual and physiological responses to Jackson Pollock's fractals. *Frontiers in Human Neuroscience*, *5*, 60. doi:10.3389/fnhum.2011.00060 PMID:21734876

Torta, E., van Heumen, J., Piunti, F., Romeo, L., & Cuijpers, R. (2015). Evaluation of unimodal and multimodal communication cues for attracting attention in human–robot interaction. *International Journal of Social Robotics*, *7*(1), 89–96. doi:10.100712369-014-0271-x

Wagatsuma, H., & Yamaguchi, Y. (2007). Neural dynamics of the cognitive map in the hippocampus. *Cognitive Neurodynamics*, *1*(2), 119–141. doi:10.100711571-006-9013-6 PMID:19003507

Wykowska, A., Anderl, C., Schubö, A., & Hommel, B. (2013). Motivation modulates visual attention: Evidence from pupillometry. *Frontiers in Psychology*, *4*, 59. doi:10.3389/fpsyg.2013.00059 PMID:23407868