A Physically Embodied Robot Teacher (PERT) as a Facilitator for Peer Learning

Bosede I. Edwards
Imagineering Institute
Iskandar Puteri, Johor Malaysia
City, University of London
United Kingdom
bosede@imagineeringinstitute.org

Idris O. Muniru
Imagineering Institute
Iskandar Puteri, Johor Malaysia
Universiti Teknologi Malaysia
Johor, Malaysia
idris@imagineeringinstitute.org

Nosiba Khougali
Imagineering Institute
Iskandar Puteri, Johor Malaysia
Universiti Teknologi Malaysia
Johor, Malaysia
nosiba@imagineeringinstitute.org

Adrian D. Cheok
Imagineering Institute
Iskandar Puteri, Johor Malaysia
City, University of London
United Kingdom
a@imagineeringinstitute.org

Rui Prada
INESC-ID and Instituto Superior Técnico, Universidade de Lisboa
Portugal
rui.prada@tecnico.ulisboa.pt

***Abstract*—This research full paper describes our work on the development, and evaluation of a physically-embodied robot teacher (PERT) as a facilitator for Peer Instruction (PI) in a high school science classroom. Embodied pedagogical agents are particularly useful in teaching young learners because of the engagement, and enjoyment they engender. Teaching robots are among possibilities in the future classroom; they may also represent solutions to the looming, acute global teacher scarcity. However, their design creates unique computational challenges. Moreover, robot-human systems involving a single robot in interaction with multiple humans, are not common. PI has been reported in many studies as a learner-centered approach that supports conceptual learning, metacognition, cognitive load sharing, and higher order learning, through peer scaffolding. However, existing studies have reported only applications by human teachers in regular classrooms. This study highlights an approach for enabling a PERT as a PI facilitator. Based on conflicting reports of the significance of agent embodiment for learning, we present the report of two experimental studies from our assessment of the system, employing the mixed method approach. We compared participants' perception of the robot's social presence and teaching quality when the robot is physically present and when it is absent (voice-only). Our findings are based on data from 14 participants in each study. The Social Presence indeX (SPiX) and the teacher quality index (TQI) were used for quantitative data collection for social presence and teaching quality, respectively. Qualitative data were based on live observation and participants' open-ended feedback. Quantitative findings show no significant difference in participants' perceptions in both conditions but qualitative feedback provided more information. Our findings show: i) students are willing to accept non-human teachers, and ii) their expectations of them are similar to those for human teachers; iii) an improved affective system will promote increased presence; iv) participants noted the robot as 'neutral', with important implications for culturally-responsive education; v) robot teachers could be potential technology-inspired solution to global teacher scarcity and vi) in addition to promoting deeper learning, PI can be a useful pedagogical approach for non-human teachers. This study is in line with anticipated ability of future technological advancements to support machines that are as effective as intelligent human tutors.**

Keywords*—*robot teacher; embodiment; pedagogical agents; PAs; peer instruction; social presence; teaching quality.

I. Introduction

Embodied pedagogical agents (PAs) can support fun, and enjoyment and can be very useful and effective with young learners, especially when engaged within collaborative learning. Teaching robots are among the possibilities envisaged in the future classroom. Key challenges with their design however, includes the ability of such systems to simulate important concepts in human-human interaction that are significant within learning environments (LEs). Robot-human systems were also mostly one-robot-one-human rather than one-robot-multiple-humans. This has educational significance especially in terms of collaboration, and requires appropriate pedagogical paradigms for effective implementation.

PI is a learner-centered, learner-led, collaborative approach, whereby students are guided through a series of well-defined steps intended to foster conceptual learning and metacognition. It incorporates peer-teaching within social learning. PI combines the advantages of many techniques in the design of its step-wise procedure, making the learner the focus of instruction, and the instructor a true facilitator. Its implementation leverages procedures deliberately intended to foster higher order thinking, and it has been reported in many studies as an effective approach for enhancing meaningful learning [1], motivation [2] and for promoting student voice, and social learning, among others. However, these studies have reported only applications by human teachers in regular classrooms. The possibility of an autonomous instructional system engaging PI approaches has not been the focus of many studies. This study highlights an approach for enabling an autonomous robot as a PI facilitator. We report on a single robot, the physically-embodied robot teacher (PERT), in interaction with multiple humans in a regular classroom setting. We also report the design challenges, responses from learners to the PERT system, and how we address affect in the design of the system. The PI approach is employed in the study, highlighting the PERT in a facilitator role in line with [3], on the ability of current technological advancements to support machines capable of playing the roles of intelligent human tutors. The key novelties of this study are focused on students' response to a non-human teacher; the role of an improved affective system in promoting increased presence,

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

students' perceptions of the robot in the teacher role and the potential implications for education, as well as the role of PI as a pedagogical approach for non-human teachers. The study aims to make further contribution to the literature on the potential of future technological advancements to support machines that are as effective as intelligent human tutors. The remaining part of this paper is divided into four sections describing related work (II), the PERT system (III), proxemics in PERT system design (IV), evaluation of the initial prototype (V) and, conclusion and future works (VI).

II. Related Work

A. Physically Embodied Pedagogical Agents

Pedagogical agents are simulated, visually represented, learner-content interfaces. Within LEs, they can serve as instructors or motivators while interacting with learners through gestures, facial expressions and natural language [1]. Several studies have reported on PAs and their educational significance [2]–[4]. The role of the PAs' image, voice, and language style as well as the significance of coordinated behaviors in its believability has also been investigated [5]. As early as the mid-80s, researchers [6] had contended that technological developments within computing, artificial intelligence (AI) and cognitive psychology had reached the point where they can support computer systems that are as effective as intelligent human tutors. 'Methodologies that will make qualitative changes in the ability to instruct students [7] however requires further investigation. In line with [7], [8] reported effect size 0.76 for Intelligent Tutoring Systems (ITS)'s effectiveness that is comparable to that of human tutors of 0.79. Their work suggests that the next-generation of ITSs may be capable of successfully playing the roles of independent instructors. Physical embodiment of PAs is in line with the theory of embodied cognition. According to [9], 'the optimal embodied intelligent agent is a robot, not distinguishable from a human'. However, human communication transcends spoken or body language, and includes stereotypes, therefore, social constructions, preconceptions and other affective concepts will still remain challenges for the best programme for embodied PAs.

B. Social Interaction and Classroom Engagement

Social interaction had focused on human relationships and leveraged verbal or non-verbal communication in all settings, including LEs. [10] emphasized the impact of social interaction on student learning, and, the benefits of social learning have been in their ability to leverage possibilities within social interaction settings. Social interaction is thus a key factor in the design of any learning system. Various studies [11], [12] have emphasized the critical role of non-verbal communication in promoting learning. Teachers' non-verbal behaviour are known to also contribute to students' judgments about their competence or character [13]. Recent changes in social interaction within human-nonhuman contexts necessitates novel theoretical ideas on people's attitudes and behaviors toward nonhuman entities. Since previous classrooms were designed for practical purposes, designs in the digital age must also address factors that are critical to students' learning and changes in the society. In the age of advanced technology-aided instruction, AI, and robotics, issues like classroom-based, human-nonhuman interaction and their significance for education, requires careful consideration.

C. Social Presence and Agency

Social presence, and agency, refer respectively to 'the sense of being there' [14] and 'the capability of individuals to make choices and act on those choices' [15]. As machines are becoming acceptable features of everyday human life, they are also seen playing many roles within LEs. Robots as teachers is still novel, and studies reporting robots as independent instructors are scarce, though the concept is already being discussed [16], [17]. A robot has no concept of personal space or awkwardness [16]; however, social presence or agency will be impossible for such a teacher. Moreover, it has been demonstrated that space-aware agents exhibit better natural behaviour [18]. Social presence is directly linked to agency. Within classroom interactions, the questions of agency, and learners' perception of teaching robots, compared with human teachers, are legitimate. In addition, the significance of agent embodiment for learning remains inconclusive in the domain of education research. To address these questions, we implemented the PERT system. We intend to test some hypotheses regarding the concept of embodiment, and how this impacts its capability as a teacher, and its social presence, based on students' perceptions. To assess social presence, we evaluated its indicators based on [14], capturing 12 indicators of social presence as physical, social and affective attributes in the Social Presence indeX (SPiX). More details are provided on instrumentation in Section V.

D. Classroom Proxemics

In the earliest conceptions of proxemics, spaces were described only in terms of distance between interactants [19]; Hall [20] ascribed specific metric values to these distances. An expanded idea of proxemics [12] however captures additional information on identity, movement, and orientation [21]. These factors, when integrated with time, can provide very useful information in interaction design and can be especially important for programming machine response. Factoring proxemics information into interaction design of a robot teacher can thus contribute towards creating a system that closely approximates natural expectations from a human teacher. In addition, mapping the orientation of relating entities, can provide information about student engagement, with potential to support the teacher in achieving classroom management [22]. Similar approaches are employed in context-aware learning systems [23], [24], and robotic systems [25]. Overall, proxemics elements have significance for classroom management and student engagement. Attention, a key precursor of engagement [26], is the inherent focus of proxemics designs, with instructional or classroom proxemics [27] focusing on the use of classroom space. Classroom proxemics can therefore represent an important factor for building the architecture for a robotic instructor.

E. *Effective Learning through PI*

The design of PAs is mostly situated within the theories of instructional scaffolding [28] which addresses the role of a 'more knowledgeable other' or MKO in providing scaffolds to assist a learner to move across a gap in level of understanding or competence, to another level. This gap is known as the 'zone of proximal development' (ZPD). Scaffolds are conceived as temporary support systems provided by the MKO, who may be an instructor, or a peer who had already crossed the ZPD. PI exemplifies this MKO role, and it is achieved through the creation of social learning opportunities within collaboration. Peer scaffolding in PI is the focus of the peer discussion sessions which also supports conceptual learning through the quiz and voting. PI is based on active learning theories [29]; and among activities considered useful for active learning are collaborative and communication activities and use of clicker questions, which are all directly linked to PI [30].

During PI, the instructor introduces a conceptual quiz item, students select their choice of answers from some options using Classroom Response System (CRS) or voting device, after which they discuss with peers on the rationale behind their choice [31]. The instructor provides explanations where necessary to clarify the answers. PI thus supports conceptual learning through an unconscious metacognitive strategy, in addition to making the class more interesting and engaging. PI also providing feedback to the instructor for pedagogical improvement. Due to the central place of the learner in the PI framework, the method can support the role of a non-human instructor within a collaborative LE. The combination of factors related to proxemics, social presence and agency, and their implications for social interaction therefore underpins our design of the context-aware, PERT system.

III. THE PERT SYSTEM

A. *Hardware and System Architecture*

Proxemics interactions are visible in the design of context-aware media players [21], mobiles [32], and large-screen ambient displays [33]. However, practical designs focusing on teaching and learning are scarce. We describe here, the design and development of the PERT system. We employed VStone's[1] so(cial) ta(lker) or Sota robot, in our work. Sota robot is shown in Figure 1. It was originally designed as a companion robot for the elderly; however, for the development of the PERT system, Sota was reconfigured as a basic science teacher. The choice of subject content being in line with current focus on STEM areas.

Sota is 24cm high, legless, with many degrees of freedom in its head, neck, eyes, and arms. It can rotate on its base and move its arms, neck, and head as well as change the colour of its eyes and lips. These features were leveraged for simulating affect as well as social presence and agency in our project. There is an internal logic sub-system that controls and coordinates the PERT system by deploying several timing controls and functions, and yielding social cues (motion, facial expression, etc.) as outputs. It is written in python language based on the availability of free programming libraries (API) and the ease of deployment in our project.

Fig 1. Sota Robot

A router coordinates data integration between sensors, the voting devices, and the robot, allowing for data transmission within the system.

B. *Instructional and Classroom Response Systems*

The instructional system is based on PI, with focus on the elements of ConcepTests, voting, and peer discussion [34]. The effectiveness of learning through PI lies in its procedure; with the discussion sessions representing the 'learning-by-teaching' process. The CRS employed was developed by our team; it comprises of a web application (to support accessibility across various mobile platforms) written in HTML and PHP and a database server. Information received through input from students' voting are stored in the database which is hosted on MySQL server from where the results are displayed as charts that can be viewed on a projector screen. The robot's motion library is provided by VStone Co. Ltd; it is written in Java language and allows manipulation of the robot's motion, and posture. A data socket summons the java programmed motion via the internal logic program written in python language. Several modules are implemented in the design for proper coordination; there were silence, text-to-speech, slide display, and quiz display modules among others. An example of the text-to-speech module is shown below:

```
def talk(statement):
    engine = pyttsx.init()
    engine.setProperty('rate', 120)
     #   voices = engine.getProperty('voices')
     engine.setProperty('voice','HKEY_LOCAL_MACHINE\SOFTWARE\Microsof
     t\Speech\Voices\Tokens\TTS_MS_EN-US_ZIRA_11.0')
    engine.say(statement)
    engine.runAndWait()
```

C. *Lesson Modules and Implementation*

Lesson modules are loaded as graphic files on the robot's instructional programme. A complete lesson is captured in a folder assigned an identification tag for the topic (e.g. CHEM 101 was the first basic chemistry lesson). The system can capture multiple subjects and/or topics in this manner. At the beginning of the class session, the PERT welcomes the students to class, introducing himself, '*Hello, my name is Sota.*

[1] https://www.vstone.co.jp/english/

I am your chemistry teacher…' An external PC augmenting its internal processing system is connected to a projector and screen for displaying information on the lesson. It then requests the students to introduce themselves, '*please mention your name and your hobby*'. The system activates the silence module at this point and waits for each student to make a personal introduction. It closes this session by relying on the same module which enables it to sense the absence of further response from students. The system is programmed to enable continuation of the session after 5 seconds of silence.

After the introduction, the PERT requests students to register their attendance using the CRS application provided. The app captures name and image input from the user. The information is deposited in a database accessible to the PERT for identifying i) students present in a specific class session, ii) students who have/have not submitted response to the quiz items, and iii) to select a random student to start the discussion session. It begins the teaching session, with '*today, we are going to be discussing the topic…*' the screen display changes, showing the first lesson slide and topic. The PERT continues the lesson, posing the first ConcepTest, displayed on the screen, and asks students to select their answer by using the CRS applications provided. This is the voting procedure in PI. Students' choices are displayed on the screen while the teacher prompts for discussion by requesting students to justify their answer choices by discussing with their colleagues.

IV. PROXEMICS IN PERT SYSTEM DESIGN

A. Login and Registration:

The initial login with name and image capture employs the GUI of the voting application; it is required only at first login. It enables the teacher to 'identify' students in the class and to call up students' names, for example, during discussions. This use of students' names, have significance for classroom management, social interaction, presence and agency and is been reported to have important implications for learning.

B. Greetings

They are considered an expression of interest; they provide a platform for building interaction, creating an atmosphere of acceptance and foundation for further interaction. [35] emphasized the place of greetings as a necessary requirement for participation in any speech community, and teacher greetings have also been reported to increase test scores [36]. All these underscore the place of greetings in social interaction. The PERT's first attempt at building a link with the students was therefore greetings.

C. Attention Allocation:

The teacher's ability to 'recognize' students' presence in the class is central to agency and social presence. Student identification by the PERT system provides a means for supporting 'teacher immediacy and social presence' [62], confirming the significance of attention allocation for higher levels of social presence or agency. Apart from being able to call out students' names, the robot could also randomly select a student to lead the peer discussion sessions.

D. Turn-taking:

Turn-taking is a rule in social interaction; it is linked to politeness and appropriateness of conduct. During the introduction and discussion sessions, the robot 'senses' silence which it interprets as time to take turn to be quiet, or continue the on-going process, thus obeying the turn-taking rule of social interaction.

V. EVALUATION OF THE INITIAL PROTOTYPE

Following, we report two studies conducted with the PERT system. Reports of studies on the significance of agent embodiment for learning is conflicting, while some researchers report no effect on learning [37], significant effects were reported for embodiment and presence in certain conditions and no effect in other conditions [38] for the same system. Hence, we set out to assess the significance of embodiment and presence for students' perception of an agent's social presence and teaching quality. Detailed objectives of each study are provided in the appropriate sections. Both studies however, have a common objective of identifying areas for system improvement. Some participants in the studies are shown with the PERT in Figure 2.

Figure 2: Some participants from a Johor Bahru High School in the study

Mixed method design was employed in both studies with data collected from 14 participants in each study. For both studies, two conditions were specified and participants were divided into two groups of 7 for the conditions. In condition 1, the robot delivered the lessons and interacted with the participants based on audio only; this is the without-robot session. In condition 2, the robot was physically present during the lesson. The same lesson was however delivered in both cases and the same procedure was maintained with both groups. Table I presents the summary of the studies and how the research objectives are addressed and presented in the mixed design study.

TABLE I: SUMMARY OF STUDY FOCUS, METHOD AND RESEARCH OBJECTIVES

Study	RQ	Focus	Quant	Qual
I	1	Rating of PERT's social presence based on embodiment type	√	
	2	Significance of embodiment for rating of SP	√	
	3	Learner expectations		√
	4	System's effectiveness		√
II	5	Rating of PERT's teaching quality based on embodiment type	√	
N/A	6	Significance of embodiment for rating of TQ	√	
All	7	Implications of findings, project challenges & improvement plans	Sum of findings	

A. STUDY 1: INITIAL ASSESSMENT OF THE PERT SYSTEM

The first study aims to assess: i) learners' perception of the robot as a socially present entity based on its physical presence or absence; ii) participants' expectations and perceptions of the teacher in terms of personality, teaching style, and limitations; iii) participants' perception of system effectiveness for promoting conceptual learning.

The quantitative instrument employed is the social presence index (SPiX), a 33-item questionnaire, based on the networked minds social presence inventory [14]. It assesses social presence based on physical (PHY), social (SOC) and affective (AFF) attributes sub-divided into 8 sub-attributes. We report the internal consistency as a measure of reliability for the instrument based on coefficient alpha. Alpha for SPiX is 0.90. However, since the co-efficient alpha is a reliability measure associated with a sample characteristics [39], we calculated alpha for the instrument. Co-efficient alpha 0.89, shows it is a reliable instrument [40] for the purpose of this study. The number of items and summary descriptions of the sub-sections of SPiX are provided in Table II:

TABLE II: DESCRIPTIONS OF THE SECTIONS AND SUB-SECTIONS OF SPIX

Attribute/ No. of Items	Sub-attributes	Description
Physical/17	PEMP	physical embodiment or presence
	PATN	attention allocation to other agents
	PMSU	perceived reciprocal message understanding
	PSYM	believable verbal/non-verbal behavior & use of symbolic expression
Social/13	SRUL	observation of the rules of social interaction
	SCTX	communication; and simulation of contex-related/specific behavior
	SMEM	social memory
Affective/3	AFF	reciprocal affective understanding, and behavioural interdependence, clear emotional system

Responses to the SPiX are based on 5-point Likert feedback with Strongly Disagree (SD), Disagree (D), Neutral (N), Agree (A) and Strongly Agree (SA).

Quantitative Results and Discussion

Research Question 1: How did learners perceived the PERT as socially present based on physical presence or absence of the robot?

Feedback is based on Likert responses to the SPiX, showing overall disagreement (SD + D), Neutrality (N), and overall agreement (SA + A). Responses for the group with a physical robot present is shown in Figure 3, and for the group with no physical robot in Figure 4. Figure 3 shows that a good percentage of the respondents in the 'without robot' group did not take a position regarding the robot's social presence as seen in the high percentage of neutral responses. For the group without robot, disagreements were almost always higher than agreement (7 of 8 cases), with higher agreement noted only for social memory (SMEM) which incidentally, is the only case where there is no disagreement.

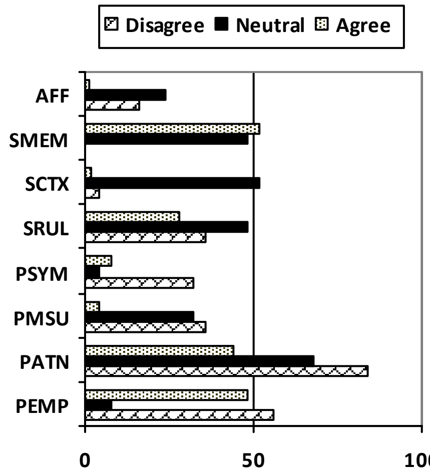

Fig. 3: Percentage Score on Social Presence Sub-Attributes by All Participants based on SPiX Responses by Participants in the 'Without Robot' Group. Items on the y-axis are based on the descriptions in Table I.

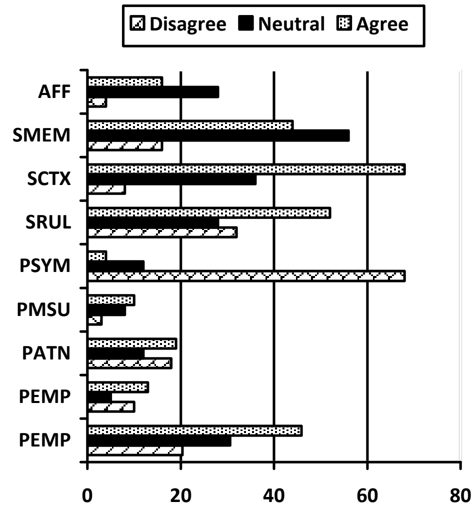

Fig. 4: Percentage Score on Social Presence Sub-Attributes by All Participants based on SPiX Responses by Participants in the 'With Robot' Group. Items on the y-axis are based on the descriptions in Table I.

This is not surprising, considering that the PERT could address participants' by name. The high number of neutral responses could not be easily explained though it could be attributed to the 'faceless' status of the 'teacher'. The high negative perception of the teacher's social presence is in line with key challenges already noted with socialization in virtual LEs [41]. This may be directly related to the absence of the socio-psychological dimension of the interaction with the physical absence of a robot since the properties of an unseen agent may not be easily evaluated. The fact that all the participants were having a novel experience with the PERT system (based on zero prior social encounter with a physical robot reported by all) is also a factor that might have affected their response.

Feedback from the group with a physical robot (Figure 4) is in contrast to that for the 'without robot' group. Agreements on the robots' social presence were higher (7 of 8 cases) than disagreements. Higher disagreements were noted for the use of symbolic expressions and non-verbal behavior (PSYM) and affective attributes (AFF) with high ratings for social attributes. The PERT's programming supports a simulated recognition of other agents and expressions indicating reciprocal perception of the same, and may explain the observation. However, since the PERT was only capable of basic programmed motions and simulated responses; the place of quick, believable and appropriate responses in addition to context-related and context-specific behaviours might explain this finding. Instructor immediacy and presence, with a linear combination of both have been reported to have statistically significant positive relationship with student motivation, cognition and affective learning [42] among others. Factors capable of improving immediacy [43] can thus contribute to presence and agency. Improved affective attributes as well as more rounded social and physical attributes can increase the believability of the non-human teachers and heighten learners' perceptions of its social presence and agency with consequent improvement in satisfaction, motivation and overall learning.

Research Question 2: How does the physical presence or absence of the robot affect students' perception of its social presence?

To answer RQ2, we evaluated the significance of the difference in rating by the two groups. The test statistics is presented in Table III. The test shows no significant difference in the overall evaluation as well as 7 of the 8 attributes by the two groups and indicate that embodiment type does not significantly influence participants' rating of social presence. However, similar to the observation in the descriptive statistics, there was a significant difference in rating for SCTX (communication and simulation of context-related or context-specific behaviour). This is easily understandable from the capabilities of the PERT.

TABLE III: TEST STATISTICS FOR COMPARISON OF SOCIAL PRESENCE

*	PEMP	PATN	PMSU	PSYM	SRUL	SMEM	SCTX	AFF	P	S	SP
Mann-Whitney U	13.0	16.50	11.0	18.00	13.00	21.00	9.00	12.50	14.50	12.00	11.50
Wilcoxon W	41.0	44.50	39.00	46.00	41.00	49.00	37.00	40.50	42.50	40.00	39.50
Z	-1.6	-1.04	-1.74	-.85	-1.48	-.45	-2.00	-1.56	-1.28	-1.61	-1.66
Asymp. Sig. (2-tailed)	0.10	0.30	0.08	0.40	0.14	0.65	0.05	0.12	0.20	0.11	0.10

*See item descriptions/legend in Table I | P=physical attributes; S = social attributes; SP = overall social presence

[38] reported more accurate perception of non-verbal communication when the embodied agent is physically present though also cautioned that the significance of embodiment may be linked to limitations of the agent.

Qualitative Results and Discussion

In this section, we address research questions 3 and 4, based on the qualitative data from observation and open-ended feedback from participants.

Research Question 3: What are learners' expectations and perceptions of the PERT in terms of personality, teaching style, and limitations?

We focus on participants' perception of the robot's social presence, evaluated based on agency, classroom management, and overall teaching and learning. We asked participants to comment on the 'personality' of the 'teacher' and their overall perception of the PERT system. Participants in the 'no robot' group noted 'something missing' in the system. They noted how a physical robot might influence the system; especially, on the significance of being able to 'see' and relate to 'someone' and his/her verbal and non-verbal expressions. Two participants described the experience of the robot's absence as similar to '*speaking to someone on the phone*'; noting that this is never the same as '*being with*' the person. This underscores the significance of non-verbal communication [12].

An important concept and unexpected point noted is what participants described as the 'neutrality' of the PERT. This is an accidental feedback that may be connected to the fact that the PERT is not racialized, gendered or assigned other common demographic characteristics [44]. This is an important issue in our multi-racial experimental context where learners and their teachers may have very different socio-demographic characteristics. The finding highlights the potential of non-human teachers for promoting reduction of bias and for culture-sensitive education in line with [45]'s observation within human resource setting. However, this finding requires further investigation.

Research Question 4: How do participants' evaluate the system's effectiveness for promoting classroom interaction and conceptual learning?

Participants found the robot's use of their names very interesting. This concept, related to physical and social attributes appear to play very important role in the perception of the robot's social presence. They showed excitement and interest, smiling and surprised when their names were randomly called by the robot. In addition, they paid more attention to the lessons, and the quiz sessions were found interesting. Participants commented on how the robot made them to review their basic chemistry and how the questions challenged them to re-evaluate their previously held conceptions. Most referred to the Q&A method as very effective for supporting memory. Others commented on the PI approach as promoting engagement. They also noted the discussion (peer learning sessions) as very good learning support. These feedback highlight the metacognitive values of PI and ConcepTests for promoting learning [46].

B. *STUDY 2: CLASSROOM EVALUATION OF PERT FOR RATING OF TEACHING QUALITY*

The second study aims to assess i) students' rating of the PERT's teaching quality; ii) the significance of embodiment for students' rating of the PERT's teaching quality; iii) the effect of the PI pedagogical framework for students' learning and the design of the PERT system.

For quantitative data collection, we employed the Teacher Quality Index (TQI), a 33-item questionnaire based on established teacher quality indices. The TQI has 7 sections as described in Table IV. Qualitative data were obtained from written feedback from participants on an open-ended protocol provided. The protocol has 4 items on which participants assessed the PERT in terms of personality, teaching method/interaction style, factors that contributed most to learning, and areas for improvement. We also asked participants to indicate prior interaction with a physical robot.

TABLE IV: DESCRIPTION OF SECTIONS OF THE TQI

Section Label	Description	No of Items
SKP	Subject Knowledge and Preparation	3
IPG	Instruction & Pedagogy	8
ASS	Assessment	4
CLM	Classroom Management	5
MAI	Motivation and Affective Input	6
FDB	Feedback	3
ENF	Engagement and Fun	4
SP	Overall Social Presence	33

Quantitative Findings and Discussion

The TQI instrument captured ratings of the PERT by participants to address RQ4 and we tested the significance of the difference of TQI for the two groups to address RQ5. We compared median values of TQI for both groups in a similar manner to the direct rating scores in Study 1 with the SPiX ratings. We report findings based on the research questions in the following sections:

RQ 5. How did participants' rate the PERT's teaching quality based on physical presence or absence of a robot?

Descriptive statistics for the 'with-robot' and 'without-robot' groups are shown in Tables V and VI. Comparison of the TQI mean values show that ratings for the group with a physical robot are always higher than for the group without a physical robot, indicating that physical presence might play important role in participants' TQI rating of the PERT.

TABLE V. DESCRIPTIVE STATISTICS FOR 'WITH ROBOT' GROUP

	SKP*	IPG*	ASS*	CLM*	MAI*	FAB*	ENF*	TQI*
Mean (μ)	10.57	30.43	14.71	16.43	20.57	8.43	14.71	115.86
Median	10.60	30.00	14.50	16.00	20.33	8.00	15.33	104.00
Std. Deviation	1.62	5.03	2.56	2.37	4.93	3.74	2.81	20.34
CV** (SD/Mean)	0.15	0.17	0.18	0.15	0.24	0.43	0.18	0.20
z-score [i.e. (μ-4)/SD]	4.06	5.25	4.18	5.24	3.36	1.18	3.81	5.50
Percentile rank (2-tail)	0.00	0.00	0.00	0.00	0.08	23.80	0.01	0.00
Percentile (left tail)	100.00	100.00	100.00	100.00	99.96	88.10	99.90	100.00
*See item descriptions in Table II; **CV: Coefficient of Variation								

The highest difference is noted for rating on 'engagement and fun'. The measures of central tendency show relative consistency in participants' attitudes. For more meaningful interpretation, the z-score-to-percentile-rank and co-efficient of variation (CV) [47] are evaluated. Z-score factors in variability in the score, it is most precise, and normalizes raw data by using a benchmark to which the mean is compared. [48] found 80% of the number of points in a scale as a right value, and recommended 4.0 for a 5-point scale (i.e. 5*80%) and 5.6 for a 7-point scale (7*80%), etc.

TABLE VI. DESCRIPTIVE STATISTICS FOR "WITHOUT ROBOT' GROUP

	SKP*	IPG*	ASS*	CLM*	MAI*	FAB*	ENF*	TQI*
Mean (μ)	8.50	26.13	11.75	13.75	18.50	7.25	11.00	96.88
Median	8.00	25.50	10.50	14.00	17.00	7.50	9.50	86.00
Std. Deviation	2.33	4.64	3.06	3.24	3.66	3.11	6.09	21.92
CV* (SD/μ)	0.27	0.18	0.26	0.24	0.20	0.43	0.55	0.23
z-score [i.e. (μ-4)/SD]	0.21	4.77	2.53	3.01	3.96	1.05	1.15	4.24
Percentile rank (2-tail)	83.37	0.00	1.14	0.26	0.01	29.37	25.01	0.00
Percentile (left tail)	58.32	100.00	99.43	99.87	100.00	85.31	87.49	100.00
Skewness	2.35	0.45	0.51	-1.25	1.39	-0.16	0.54	0.83
*See items' legend/descriptions in Table II; **CV: Coefficient of Variation								

We used the 4.0 benchmark and the z-score is calculated from the ratio of difference of mean and benchmark to SD. Higher CVs connote higher variability and noticeably different CVs indicate inconsistent respondents' attitudes [47].

Further analysis of the descriptive statistics shows relatively consistent participants' attitudes (CVs) for 'with-robot' group. All scores fall only 0.15-0.24 SD below the benchmark. A slight variation is noted with the feedback (FDB) with a slightly higher CV (0.43) and ratings falling 0.43 SD below the benchmark. This is also noted in the lower z-score compared with values for other indices. Higher variability and less consistent participant attitudes are noted for the 'without-robot' group. The highest CV values are noted for FDB and ENF with scores falling 0.43 and 0.55 SD below the benchmark. The z-score values show a different pattern compared with the 'with robot' group while percentile values (left tail) follow the same trend.

RQ6. Does the physical presence or absence of a robot have significant effect on students' assessment of the PERT's teaching quality?

To answer this question, we tested the null hypothesis:
H0: There is no significant difference in participants' rating of the robot's TQ based on its physical presence or absence.

We computed the mean difference of TQI ratings for presence and absence of physical robot for the two groups. Details are presented in Table VII.

TABLE VII. MEAN TQI RATINGS OF THE PERT'S TEACHING QUALITY

	SKP*	IPG*	ASS*	CLM*	MAI*	FAB*	ENF*	TQI*
Mean (with-robot)	10.57	30.43	14.71	16.43	20.57	8.43	14.71	115.86
Mean (without-robot)	8.50	26.13	11.75	13.75	18.50	7.25	11.00	96.88
Mean Difference	2.07	4.30	2.96	2.68	2.07	1.18	3.71	18.98
*See item descriptions in Table II; **CV: Coefficient of Variation								

For the small sample size (n<2000), we focus on non-parametric statistical analysis of the data, the test statistics are shown in Table VIII.

TABLE VIII. MWW TEST FOR MEAN DIFFERENCE OF TQIS

	SKP*	IPG*	ASS*	CLM*	MAI*	FAB*	ENF*	TQI*
Mann-Whitney U	8.50	13.50	12.50	14.00	20.50	22.00	16.50	13.00
Wilcoxon W	36.50	41.50	40.50	42.00	48.50	50.00	44.50	41.00
Z	-2.08	-1.41	-1.55	-1.38	-.52	-.32	-1.03	-1.47
Asymp. Sig. (2-tailed)	.04	.16	.12	.17	.61	.75	.31	.14

Though the statistics show that participants' in the 'with robot' group rated the robot higher than those in the 'without robot' group, the difference is not significant except for SKP ($p<0.05$). The finding suggests that overall, ratings are not influenced by the robot's physical presence or absence. There may be a link in the noted significant difference in SKP (subject knowledge) and the influence of teacher's non-verbal communication on students' perception of a teacher's competence and knowledge noted earlier [13].

C. IMPLICATIONS OF FINDINGS, PROJECT CHALLENGES, AND IMPROVEMENT PLANS

In this section, we address the last research objective which captures implications of our findings for the future classroom, brief description of challenges with the PERT system development and improvement plans for the system.

i. Implications of findings for the future classroom with robots as teachers

Qualitative responses from participants show that physical and affective attributes are very central to perception of the robot as a socially present agent. Considering the factors captured in the physical attributes, it is not surprising that despite the touted effectiveness of many 'faceless' PAs, learners still consider the 'physical' presence of a 'teacher' as important. This finding highlights the significance of blended approaches and, the role of face-to-face, teacher-student interaction and teachers' non-verbal communication. The challenges of early attrition, non-completion, and dropouts in strictly virtual tutoring platforms [49] may be strongly connected to the negative impact of an impersonal instructor. Students want to 'see' the physical form of an instructor.

Though, tone of voice can contribute to the conveyance of affect, non-verbal communication cues are easily observed physically and are thus, directly linked to physical presence. This suggests a correlation between physical and affective attributes. Some participants noted that the PERT used 'harsh' expressions; commenting that the PERT sounded 'rude' and authoritative, and citing its use of phrases like '*you have 40 seconds to do that*', and '*do that in the next 20 seconds*' etc. However, we noted that this and it may be connected to socio-cultural factors which is outside the scope of the current study.

ii. Challenges with the initial system and plans for system improvement?

In the initial system, we used fixed-time programming for the turn-taking. The lengths of discussion sessions however vary based on the number of participants, making this troublesome. We implemented the silence detection module to address this challenge. However, in noisy environments (which are better approximates of real-life LEs) we had to increase the sensitivity of the system which in turn became overly sensitive at lower noise levels. This underscores the significance of responsiveness in future programming of such systems. Another issue observed was the difficulty with accessing the voting application which is not fully deployed as a mobile application yet, therefore requiring connection through a local network router. This necessitates connection using a specific IP address. To address this, we provided pre-configured mobile phones for participants during the study sessions, but we intend making the CRS a full mobile app during future upgrades. We also noted that when answers were displayed after voting, most participants were unable to remember details of the answer choices provided earlier. Hence, we modified the display to show both the answer choices and responses side-by-side.

VI. CONCLUSION AND FUTURE WORKS

Growth in AI and robotics has great implications for the future classroom. Despite assumptions by many to the contrary, the emergence of full-fledged robot teachers appears inevitable. Current teacher scarcity in many subject areas and scarcity of highly skilled teachers, in addition to the economic advantages of machines over humans further support the speedy progress of AIED in this direction. We implemented a robot teacher employing the peer-teaching approach with focus on the benefits of social and conceptual learning in high school science. Overall, our design supports the simulation of social and emotional intelligence in classroom-based HRI. Additional novelties of our work include the development of a basic architecture for enabling the design of realistic robot teachers and the significance of a robot's 'neutrality' for fostering inclusiveness and cultural responsiveness in 21st-centry classrooms. While recognizing the limitations of the current system in terms of hardware, software and scope, it represents a basis for future development. The importance of user creativity in the design of a learning system cannot be overemphasized, hence, future work should consider user-friendly, teacher-operable systems and those that employ reusable templates based on GUI. Cost issues also needs to be addressed. Finally, the concept of neutrality noted by participants also requires further investigation.

ACKNOWLEDGMENT

This work was partially supported by funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013

REFERENCES

[1] G. Veletsianos and G. S. Russell, "Pedagogical agents," in *Handbook of research on educational communications and technology*, Springer, 2014, pp. 759–769.

[2] W. L. Johnson and J. C. Lester, "Face-to-Face Interaction with Pedagogical Agents, Twenty Years Later," *Int J Artif Intell Educ*, vol. 26, pp. 25–36, 2016.

[3] G. Veletsianos and G. S. Russell, "Pedagogical Agents," *Handb. Res. Educ. Commun. Technol.*, pp. 759–769, 2014.

[4] R. Moreno, R. E. Mayer, and J. C. Lester, "Life-Like Pedagogical Agents in Constructivist Multimedia Environments: Cognitive Consequences of their Interaction," *Proc. World Conf. Educ. Multimedia, Hypermedia Telecommun. 2000*, no. 1997, pp. 776–781, 2000.

[5] J. C. Lester, J. L. Voerman, S. G. Towns, and C. B. Callaway, "Deictic believability: coordinated gesture, locomotion, and speech in lifelike pedagogical agents," *Appl. Artif. Intell.*, vol. 13, no. 4–5, pp. 383–414, 1999.

[6] J. R. Anderson, C. F. Boyle, and B. J. Reiser, "Intelligent tutoring systems," *Science (80-.).*, vol. 228, no. 4698, pp. 456–462, 1985.

[7] J. R. Anderson, C. F. Boyle, and B. J. Reiser, "Intelligent tutoring systems," *Science (80-.).*, vol. 228, no. 4698, pp. 456–462, 1985.

[8] K. vanLehn, "The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems," *Educational Psychologist*, vol. 46, no. 4. pp. 197–221, 2011.

[9] M. Haake, "Embodied Pedagogical Agents: From Visual Impact to Pedagogical Implications," 2009.

[10] B. Hurst, R. Wallace, and S. B. Nixon, "The Impact of Social Interaction on Student Learning," *Read. Horizons*, vol. 52, no. 4, pp. 375–398, 2013.

[11] C. P. Zeki, "The importance of non-verbal communication in classroom management," *Procedia - Soc. Behav. Sci.*, vol. 1, no. 1, pp. 1443–1449, 2009.

[12] A. Hans and E. Hans, "Kinesics, Haptics and Proxemics: Aspects of Non -Verbal Communication," *IOSR J. Humanit. Soc. Sci. Ver. IV*, vol. 20, no. 2, pp. 47–52, 2015.

[13] A. Georgakopoulos and L. K. Guerrero, "Student perceptions of teachers' nonverbal and verbal communication: A comparison of best and worst professors across six cultures," *Int. Educ. Stud.*, vol. 3, no. 2, pp. 3–17, 2010.

[14] F. Biocca, C. Harms, and J. Gregg, "The Networked Minds Measure of Social Presence : Pilot Test of the Factor Structure and Concurrent Validity Co-Presence," in *4th annual International Workshop on Presence*, 2001, vol. 21, no. January, pp. 9–11.

[15] R. Lindgren and R. Mcdaniel, "Transforming Online Learning through Narrative and Student Agency," *Educ. Technol. Soc.*, vol. 15, no. 4, pp. 344–355, 2012.

[16] M. Powell, "Robot Teachers in the Classroom," *iQ*, 2014. [Online]. Available: https://iq.intel.com/robot-teachers-in-the-classroom/. [Accessed: 12-Mar-2017].

[17] A. J. C. Sharkey, "Should we welcome robot teachers?," *Ethics Inf. Technol.*, 2016.

[18] H. Laga and T. Amaoka, "Modeling the spatial behavior of virtual agents in groups for non-verbal communication in virtual worlds," in *Proceedings of the 3rd International Universal Communication Symposium on - IUCS '09*, 2009, vol. 5, p. 154.

[19] I. J. Arias, "Proxemics in the ESL Classroom," *Forum Qual. Soc. Res.*, vol. 34, no. 1, p. 32, 1996.

[20] E. T. Hall, "Proxemics," *Curr. Anthropol.*, vol. 9, no. 2/3, p. 83, 1968.

[21] T. Ballendat, N. Marquardt, and S. Greenberg, "Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment," *ACM Int. Conf. Interact. Tabletops Surfaces - ITS '10*, p. 121, 2010.

[22] E. T. Emmer and L. M. Stough, "Classroom Management: A Critical Part of Educational Psychology, With Implications for Teacher Education," *Educ. Psychol.*, vol. 36, no. 2, pp. 103–112, 2001.

[23] Y. Yang *et al.*, "CCESK: A Chinese Character Educational System Based on Kinect," *IEEE Trans. Learn. Technol.*, vol. PP, no. 99, 2017.

[24] T. Hsu, C. Chiou, J. C. R. Tseng, and G. Hwang, "Development and Evaluation of an Active Learning Support System for Context-Aware Ubiquitous Learning," *Learn. Technol. IEEE Trans. Learn. Technol.*, vol. 9, no. 1, pp. 37–45, 2016.

[25] W. Burlson *et al.*, "Active Learning Environments with Robotic Tangibles: Children's Physical and Virtual Spatial Programming Experiences," *IEEE Trans. Learn. Technol.*, vol. PP, no. 99, 2017.

[26] D. Lagun and M. Lalmas, "Understanding User Attention and Engagement in Online News Reading," *Proc. Ninth ACM Int. Conf. Web Search Data Min. - WSDM '16*, pp. 113–122, 2016.

[27] J. A. McArthur III, "Instructional Proxemics : Creating a place for space in instructional communication discourse," Clemson University, 2008.

[28] L. S. Vygotsky, "Interaction between learning and development," *Mind and Society*. pp. 79–91, 1978.

[29] E. Mazur, "Peer Instruction (Eric Mazur)," *Youtube*, 2012. [Online]. Available: https://www.youtube.com/watch?v=Ay3SVyDNTrg#t=21. [Accessed: 25-Jan-2016].

[30] Z. Yu, W. Chen, Y. Kong, X. Sun, and J. Zheng, "The Impact of Clickers Instruction on Cognitive Loads and Listening and Speaking Skills in College English Class," *PLoS One*, vol. 9, no. 9, pp. 1–8, 2014.

[31] B. Edwards, B. Aris, N. Shukor, and H. Mohammed, "Systematic Review of the Literature on Cognitive Load Management through Collaborative Learning Techniques: A Case for Peer Instruction." 2016.

[32] D. Ledo, S. Greenberg, N. Marquardt, and S. Boring, "Proxemic-Aware Controls: Designing Remote Controls for Ubiquitous Computing Ecologies," *Proc. of MobileCHI*, pp. 187–198, 2015.

[33] S. Greenberg, S. Boring, J. Vermeulen, and J. Dostal, "Dark patterns in proxemic interactions: A critical perspective," *2014 Conf. Des. Interact. Syst.*, pp. 523–532, 2014.

[34] B. I. Edwards, B. Aris, N. A. Shukor, and H. Mohammed, "Using Response System through Voting in Peer Instruction for Learning Sustainability," *J. Teknol.*, vol. 77, no. 13, pp. 147–157, Nov. 2015.

[35] A. Duranti, "Universal and culture-specific properties of greetings," *J. Linguist. Anthropol.*, vol. 7, no. 1, pp. 63–97, 1997.

[36] L. Weinstein, A. Laverghetta, R. Alexander, and M. Stewart, "Teacher Greetings Increase College Students' Test Scores.," *Coll. Stud. J.*, vol. 43, no. 2, pp. 452–453, 2009.

[37] J. Ryu and F. Ke, "Increasing Persona Effects : Does It Matter the Voice and Appearance of Animated Pedagogical Agent *," *Educ. Technol. Int.*, vol. 19, no. 1, pp. 61–91, 2018.

[38] A. Mollahosseini, H. Abdollahi, T. D. Sweeny, R. Cole, and M. H. Mahoor, "Role of embodiment and presence in human perception of robots' facial cues," *Int. J. Hum. Comput. Stud.*, vol. 116, pp. 25–39, 2018.

[39] I. Ercan, B. Yazici, D. Sigirli, B. Ediz, and I. Kan, "Examining Cronbach Alpha, Theta, Omega Reliability Coefficients According to Sample Size," *J. Mod. Appl. Stat. Methods*, vol. 6, no. 1, pp. 291–303, 2007.

[40] M. Tavakol and R. Dennick, "Making sense of Cronbach ' s alpha," *Int. J. Med. Educ.*, pp. 53–55, 2011.

[41] C. Irwin and Z. Berge, "Socialization in the Online Classroom," vol. 9, no. 1, pp. 1–7, 2006.

[42] C. Baker, "The impact of instructor immediacy and presence for online student affective learning, cognition, and motivation.," *J. Educ. Online*, vol. 7, no. 1, pp. 1–30, 2010.

[43] J. B. Arbaugh, "Sage, guide, both, or even more? An examination of instructor activity in online MBA courses," *Comput. Educ.*, vol. 55, no. 3, pp. 1234–1244, 2010.

[44] K. Tyson, *Integration Interrupted: Tracking, Black Students, and Acting White after Brown*. 2012.

[45] Forbes Human Resources Council, "How Artificial Intelligence Can Eliminate Bias and Improve HR Operations," *Forbes CommunityVoice*, Jersy City, New Jersey, Oct-2017.

[46] A. Zohar and S. Barzilai, "A review of research on metacognition in science education: current and future directions," *Stud. Sci. Educ.*, vol. 49, no. 2, pp. 121–169, 2013.

[47] J. Sauro, "How To Interpret Survey Responses: 5 Techniques," *MeasuringU*, 2011. .

[48] J. Nielsen and J. Levy, "Measuring usability: preference vs. performance," *Communications of the ACM*, vol. 37, no. 4. pp. 66–75, 1994.

[49] K. Tyler-Smith, "Early attrition among first time eLearners: A review of factors that contribute to drop-out, withdrawal and non-completion rates of adult learners undertaking eLearning programmes," *J. Online Learn. Teach.*, vol. 2, no. 2, pp. 73–85, 2006.