Scaffolding of Motivation in Learning using a Social Robot

Andrea Deublein, Anne Pfeifer, Katinka Merbach, Katharina Bruckner,

Christoph Mengelkamp, Birgit Lugrin

University of Wuerzburg

This is a preprint.
Full article published in Computers & Education
Volume 125, October 2018, Pages 182-190

Please find the final version here: https://doi.org/10.1016/j.compedu.2018.06.015

1. Introduction

In the societal change from an industrialized society towards a knowledge society, lifelong learning is gaining importance constantly (Gara, 2001). Knowledge as a factor of production as well as a factor of location influences the ability to compete internationally, which is essential in times of globalization (Mandl & Krause, 2001). Lifelong learning is only successful if learners are equipped with effective strategies and tools. Hence it is inevitable to explore how to optimize the learning processes and support learners in the best possible way. One key factor in this context is the motivation of learners (Roesler, 2011). Motivation can be defined as the orientation of thoughts and actions towards a positive goal (Rheinberg, 2004, p. 15). In the context of learning, motivation describes the intention to acquire skills or knowledge about a knowledge domain (Deimann, 2002), and motivation is a critical success factor as it helps to master difficult tasks and work with greater persistency (Zimmerman, 2011). The use of multimedia environments to support lifelong learning bears great potential, but is also demanding for the learners. Amongst other things, learners have to motivate themselves to start learning and not to quit the learning process. The novelty of the situation and the curiosity piqued by the technical learning support are not sufficient to keep learners motivated for extended periods of time (Keller & Suzuki, 2004). The importance of considering motivational factors while designing multimedia learning environments has already been demonstrated (e.g. ARCS model, Keller & Kopp, 1987). Especially when learners learn by themselves, without the support of a teacher or a peer, effective learning strategies help to maintain the level of motivation and thereby increase the learning outcome.

Social robots have attracted increasing attention in the last decade. They are supposed to support humans in areas such as education, therapy or care of the elderly (e.g. Admoni &

Scassellati, 2014; Mubin, Stevens, Shahid, Mahmud & Dong, 2013). In the domain of teaching and learning, social robots can be employed to assist learners during the learning process.

To explore the potential of a social robot to increase the motivation of a learner, this study combines findings from educational psychology (the ARCS model, Keller & Kopp, 1987) with the potential of a social robot to assist learners. Therefore, a learning environment was developed in which learners are supported by a social robot in learning a language. To the best of our knowledge, the potential of the ARCS model has not been explored using a social robot so far. The model provides strategies to promote the motivation of learners in a systematic way and has been used in various technology-based learning environments (e.g. Cook, Beckman, Thomas, & Thompson, 2009; Kim & Keller, 2008; Su & Cheng, 2015). We therefore used it as a basis for modelling the behaviour of the robot. The aim of this contribution is to investigate whether a social robot that exhibits motivational behaviour according to the ARCS model has an effect on the motivation of learners and, as a consequence, on the learning outcome.

2. Related Work and Theoretical Background

2.1. Social robots in the context of learning

A robot is defined as a social robot if it interacts with humans, taking into account social norms and emotions (Breazeal, 2003). One of their major fields of application is in the domain of education. Here, they are predominantly used for language learning (e.g. Saerbeck et al., 2010; Takana & Matsuzoe, 2012; Moriguchi et al., 2011) or the acquisition of technical and mathematical knowledge (e.g. Kennedy et al., 2015; Leyzberg et al., 2012).

Social robots appear to have an advantage compared to more traditional devices used in technology-based learning. Leyzberg et al. (2012), for example, have shown that a greater learning outcome can be achieved if a learning content is taught by a social robot, compared to on-screen display. In comparison to a human teacher, it has been shown that feelings of shame and anxiety can be reduced when using a social robot (Yang & Chen, 2007).

A major advantage of social robots is their ability to interact with humans in a natural and social manner. The way the robot presents social behaviour usually varies with the role of the robot, the learning material, and the age of the learner (Mubin et al., 2013). Previous research revealed evidence that the efficiency of the learning process can be increased by a social robot that behaves in a supportive way. Saerbeck et al. (2010) showed that learning a language with a social robot that acts in a socially supportive manner leads to an increased learning outcome compared to a social robot that acts in a neutral way. In addition, the authors showed that in a language learning scenario a social robot interacting in the role of a tutor was preferred. Further, Ushida (2010) documented that a robot exhibiting social behaviour can positively influence the retention of English vocabulary. Moreover, the author showed that the assistance of a social robot can also increase interest in the subject. Stressing the importance of the motivation of the learner, Kanda and Ishiguro (2005) found that the motivation of children who learned with robots increased significantly during the learning process. However, this effect was only maintained for a short period of time. Likewise, Chang, Lee, Chao, Wang, and Chen (2010) increased the motivation of primary school children by creating an interactive and hands-on learning experience with a social robot.

The majority of studies that successfully test teaching with social robots therefore focus on educating children, or children with special needs (e.g. Robins et al., 2005). However, there is a growing interest in teaching adult learners using social robots. Examples include research by Kidd and Breazeal (2008) or Schodde et al. (2017), whose results suggest positive effects from using social robots in adult learning.

In this study, we use a social robot to increase the motivation and thereby the learning outcome of learners. We therefore integrate the ARCS model in the robot's behaviour in a systematic manner.

2.2. ARCS Model

The ARCS model (Keller & Kopp, 1987) was developed from the Motivational Design Theory of Keller (1983). It provides a taxonomy of strategies that can be used to improve the motivation to learn. The ARCS model offers a set of categories and subcategories from which strategies can be selected. All the strategies are based on different concepts and theories in motivation (Song & Keller, 2001). "ARCS" is an acronym based on four motivational dimensions:

- Attention: attention and interest of the learner should be activated and maintained. Factors such as curiosity and surprising stimuli play an important role. The subcategories are: perceptual arousal (how to capture the learner's interest), inquiry arousal (how to stimulate curiosity and an attitude to ask questions) and variability (how to maintain the learner's attention by offering different methods) (Keller, 1983; Small, 1999; Visser & Keller, 1990).
- Relevance: learning goals should be relevant to the learner. A precisely defined learning goal is therefore of great importance. A link between the needs of the learners and the teaching content should be established. The subcategories are: goal orientation (how to meet learners' needs and present the objectives and usefulness of the content for the learner), motive matching (matching objectives

to the student's individual needs and motives) and familiarity (how to tie instructions to learners' experiences) (Keller, 1983; Small, 1999; Visser & Keller, 1990).

- Confidence: learners should expect a positive learning outcome and attribute their success to their own abilities. They also should experience a feeling of control and competence. The subcategories are: clarify learning requirements (how to support the structure of a positive expectation of success), create opportunities for success (support and promote the perception of one's own competence) and control option (how to help learners to use their efforts and abilities as a cause of success) (Keller, 1983; Small, 1999; Visser & Keller, 1990).
- Satisfaction: learning should trigger positive emotions. Thus it is important to provide feedback to the learner and offer the possibility of assessing their own performance. The subcategories are: natural consequences (provide the learner with meaningful opportunities to apply the newly-acquired knowledge), positive outcome (how to maintain the desired behaviour of learners) and coherence (how to help learners to evaluate their own performance as positive) (Keller, 1983; Small, 1999; Visser & Keller, 1990).

The model is based on the interplay between motivational concepts and problem-solving approaches in the design of the learning content (Keller & Suzuki, 2004).

3. Learning Environment

To investigate whether a social robot has a motivational effect when using the ARCS model strategies, a learning environment with various components was implemented: a set of questions and answers were implemented in HTML, to be shown on a screen and answered by the users. A robot that interacts in a motivational manner was set up next to the screen to interact with the user, depending on the state of his or her progress within the learning environment. Figure 1 shows the setup of our learning environment.

A Reeti robot[1] was chosen as the social robot. Reeti is an expressive social robot, produced by the French company Robopec. The robot is 44cm tall and has 14 degrees of freedom in the head and neck area. For speech output, the text-to-speech function was used. Visual SceneMaker (Gebhard, Mehlmann, & Kipp, 2011) was employed to model its behaviour, and connected to an HTML environment which the learners used to answer questions.

3.1. Learning material

Our learning environment was set up in the domain of second language learning. The language Spanish was chosen, since this is a very popular language chosen by many students at our university's Centre for Language Learning. In cooperation with the teachers of those Spanish courses, exercises were developed for our learning environment that can be implemented for digital presentation and resemble the task assignments for students taking a beginners' Spanish course (language level A1). The exercises were divided into six sections, each containing four tasks. The sections covered vocabulary, grammar and reading comprehension, and contained the content from the language course. The content of the learning material was therefore intended to be relevant to students of these classes and of a reasonable degree of difficulty.

The tasks were designed as single-choice questions with four answers to choose from. The only exception was one topic that did not allow more than two meaningful answers. This

[1] Reeti robot: http://www.robopec.com/en/products/reeti-robopec/

section therefore contained tasks with two options only. An example of a grammar task is: "¿Quién de vosotros toca… a) al fútbol?, b) algún instrumento?, c) gimnasia? d) a las cartas?" (This roughly translates as "Who of you plays a) football? b) an instrument? c) sports? d) cards?"). Please note that, according the Spanish grammar rules, just one of the answers is grammatically correct. The participants had to choose an answer, and got feedback from the robot whether the answer was correct or not. If participants had chosen the wrong answer, the robot additionally told them which of the remaining answers was correct. The statement provided by the robot was randomly chosen from nine feedback options (e.g. "That's the correct answer.", "You're right." or "That wasn't the right answer.", "No, you're wrong."). For the learning environment, we re-named Reeti into "Reyes" to better fit the learning content. This gender-neutral name was chosen to minimize gender effects.

3.2. Implementation of the ARCS Model

The ARCS model (see Section 2) had to be adapted in a way to make it suitable for the interaction with a social robot. To implement the robot's behaviour, the dimensions attention (A) and relevance (R) and likewise the dimensions confidence (C) and satisfaction (S) have been merged to two conditions. To this end, the robot's motivational behaviour was implemented either before the learning task, in order to enhance attention on the task and to promote the perceived relevance of the task (AR), or after the learning task, to increase the confidence and satisfaction of the learner (CS). In order to adapt the ARCS model on our design, we first collected different examples for each category and subcategory from Keller (1983), Keller and Kopp (1987), Kim and Keller (2008), Small (1999), Song and Keller (2001) and Visser and Keller (1990). Based on these examples we formulated statements that suited our context.

To raise the learner's motivation before the learning task, introductory statements were formulated for each task to heighten attention and relevance. In order to increase the learner's attention, it is important to arouse the learner's interest, make the learner curious and try to maintain attention by varying the methods (e.g. Keller, 1983; Small, 1999). To do so, the robot asked questions prior to the task which required the learners to think about the topic (Song & Keller, 2001). Relevance-related statements should evoke a feeling of familiarity (Kim & Keller, 2008; Song & Keller, 2001), for example by tying the learning content to previous experiences or a personal way of communication. In addition, they point out the importance or utility of the lesson (Song & Keller, 2001) and how the learned content may relate to future activities (Keller & Kopp, 1987). The following statements exemplify how attention and relevance were addressed by the robot in our study:

AR example 1: "Imagine yourself on holiday in Spain where you meet new friends. Wouldn't it be nice if you were able to talk to them about shared interests? How about hobbies, for instance?"

AR example 2: "As you already know from past lessons, Spanish verbs have to be conjugated, which means there are different cases for different people. Confusing, isn't it? Do you know which verb belongs to whom?"

In order to increase confidence and satisfaction, different variations of statements that were presented after the learning tasks were formulated. These were given in addition to the generic feedback (answer was correct/incorrect) after each task. To increase the learners' confidence, the robot emphasizes that the learners' success is due to their own efforts and skills and encourages the learners to believe that they are able to achieve their goals (Kim & Keller, 2008). Another possible strategy is to emphasize that the learners' knowledge from previous tasks will help to solve the current one (Song & Keller, 2001). The learners' satisfaction can be increased by using positive feedback as a reward or showing possibilities to apply new knowledge in everyday life (Small, 1999). The following examples show how the robot addresses confidence and relevance if the answer given by the user was incorrect or correct respectively.

CS example 1: "If you ponder over the next task a little bit harder, you will be able to solve it for sure."

CS example 2: "Well done. I'm sure you will be able to use your newly-acquired knowledge very soon."

To integrate the ARCS model into the non-verbal behaviour of the robot, we extracted behaviours from the literature that are known to have a motivational effect, and transferred these behaviours to the social robot. Examples of such behaviour are eye movements, head nods and shakes (Brown, Kerwin & Howard, 2013; Saerbeck et. al, 2010), facial expressions (Baylor & Kim, 2009), and the display of empathy (Baylor & Kim, 2005; Mori, Prendinger & Ishizuka, 2003; Saerbeck et. al, 2010). The wiggle of the robot's ears is also supposed to have a positive effect on motivation (Breazeal, 1998). Different head movements (shaking and nodding the head), facial expressions (happy, sad, surprised, confused), ear movements (wiggling the ears, tilting the ears forwards or backwards) and eye movements (blinking) were modelled using the tools R-Pilot and R-ShowMaker that are supplied with the Reeti robot. Lighting up the robot's cheeks was used as an eye-catcher.

The robot's non-verbal behaviour was integrated to fit its verbal behaviour. In the AR condition wiggling of the ears, sound or light effects, together with curious and confused facial expressions were used to attract and increase attention. Nods or shakes of the head as a reaction to the way the learner solved the task, as well as happy or sad facial expressions, were used in the CS condition of the experiment.

4. User Study

A quantitative experiment was carried out to compare the two conditions implemented (AR vs. CS) and to find out whether the social robot exhibiting motivational behaviour according to the ARCS model has an effect on the motivation of the learners and hence the learning outcome. A between-subjects design was chosen and participants were randomly assigned to one of the two conditions. The robot's motivational behaviour *before* the tasks was intended to increase attention and relevance in the AR condition, whereas the robot's motivational behaviour *after* the tasks should lead to increased confidence and satisfaction in the CS condition. More precisely, we expect that:

H1: Motivating utterances and non-verbal behaviour prior to presentation of the learning task increase the learners' attention as well as the perceived relevance of the task.

H2: Motivating utterances and non-verbal behaviour after presentation of the learning task increase the learners' confidence and satisfaction.

H3: The robot's behaviour influences the learning success mediated by the learners' motivation.

In addition, we were interested in how the robot and the interactive learning environment were perceived by the users.

4.1. Questionnaires and knowledge test

A socio-demographic questionnaire was used to ask questions about gender, age, major subject of study, and the total number of semesters the students had completed. It also recorded whether students were taking the language course voluntarily or if it was mandatory for their study, how often they study using e-learning programs, how confident they were in their own computer skills, whether they interacted with robots before in general, and with the Reeti robot in particular.

Cognitive load was measured by the degree to which participants agreed with two statements concerning how much effort the learners invested (cf., Ayres 2006), and how difficult they found the material (cf., Paas, 1992). Scales ranged from "1" = hardly agree to "100" = strongly agree. As we had no research questions relating to cognitive load, no results are presented in this paper. The degree to which the robot was perceived as being **motivating** was determined using Keller's Instructional Materials Motivation Scale (IMMS) (Huang & Hew, 2016; Keller, 1987), which uses six-point Likert Scales ranging from "I don't agree at all" to "I fully agree". The English scale was translated to German and adapted to fit the experimental design, i.e. items were re-formulated to refer to the robot. Two of the items did not fit the context of the experiment and were omitted for that reason. Finally, 34 items remained which were split between the four sub-scales: attention, relevance, confidence and satisfaction. The **enjoyableness and usefulness** of the interaction with the robot were measured using adjectives proposed by Fasola and Mataric (2012). Adjectives were translated into German and rated on six-point Likert Scales ("I don't agree at all" to "I fully agree"). To measure the perceived **intelligence** of the robot, the German translation of the relevant sub-scale from The Godspeed Questionnaire Series by Bartneck et al. (2009) was employed. Five pairs of adjectives were rated on a six-point polarity profile. The **social presence** of the robot was measured on a translated version of a scale based on Jung and Lee (2004). To rate perceived **ability** as a learning companion, five items by Fasola and Mataric (2012) were translated and adapted. We asked whether the participants felt **distracted** from their learning task by the robot ("very strongly", "strongly", "hardly", "not at all"). Finally, participants were given the opportunity to give comments on the learning experience in an open text field. The number of items per sub-scale, as well as the reliabilities of the scales, are presented in Table 1.

The **knowledge test** consisted of eleven open questions that were designed in cooperation with the teachers of the Spanish courses at our University. Eight were single-choice questions, and there was one question in which the participants had to assign 13 different words to four categories. Five of the six sections that the participants practised with the robot were covered by the language test. One section was left out as its tasks had only two answers to choose from, and thus the probability that the right answer could be chosen by chance would have been too high.

4.2. Procedure

Participants were greeted by the experimenter, signed a consent form and were randomly assigned to one of the two conditions. Each session lasted approximately 30 minutes and consisted of the following steps:

1. socio-demographics questionnaire,
2. interaction with the learning environment,
3. questionnaires (cognitive load, motivation, enjoyableness and usefulness of the interaction, social presence, perceived intelligence, ability as a learning companion),
4. knowledge test.

To emphasise that the questionnaires and the knowledge test were not part of the interaction with the learning environment, the interactive learning took place in a different workspace. The interaction with the learning environment (step 2) lasted approximately 10 minutes. Depending on the condition participants were assigned to, they were either motivated by the robot prior to each task or after each task they had to solve. In total, all participants completed six sessions containing four tasks each. The procedure for such a session is outlined in Figure 2.

4.3. Participants

A total of $N = 39$ participants took part in the study. At the time of the experiment all of them were enrolled in a basic Spanish language course (A 1) at our University. Participants were recruited during their Spanish language courses by the teacher requesting students to take part in the experiment. All participants except one were students. $N = 19$ participants were assigned to the CS condition, $N = 20$ participants were assigned to the AR condition. The participants were $M = 22.28$ years old ($SD = 2.67$) and the majority (28.20%) was enrolled in their third semester ($M = 4.72$, minimum = 1, maximum = 12). $N = 15$ (38%) were male, $N = 24$ (62%) female; genders were distributed evenly between the two conditions. None of the participants had ever been in contact with the robot Reeti before, and only two of them had interacted with robots before. While most of the participants considered themselves very skilled (33.3%) or skilled (59.0%) in handling computers, only 7.7% considered themselves less skilled or even insecure. 61.5% said they never used e-learning programs, 7.7% use them approximately once per semester, 12.8% once per week or once per month, and 5.1% more often than once a week.

Nine of the participants stated that they were not taking the language course voluntarily (but had to take it as part of their studies). The possibility of making open comments concerning the experiment revealed that, in fact, three participants had not covered the last lesson of the experiment in their group of the language course. Since those three were distributed between the two conditions, this will not be taken into consideration. Participants were paid €5 for taking part in the study.

5. Results

The main aim of the user study is to explore whether different robot behaviour can influence the learners' motivation and learning success. To answer this question, measurements were made of participants' attention, perceived relevance of the content of the lessons, learners' confidence and satisfaction. All results were compiled using IBM SPSS Statistics.

The complete descriptive values of the ARCS scales and the learning outcome can be seen in Table 2. The means of the scales for attention, confidence, relevance and satisfaction after learning with Reeti ranged from M=3.92 to M=4.64. The learning outcome in the AR condition is M = 62.81% correct answers (SD = 17.08), and in the CS condition M = 66.78% (SD = 12.90) of the answers given were correct. Therefore, participants performed substantially better than guessing, but there is room for improvement (i.e., no ceiling effects).

For all inferential analyses the Type I error rate was set to .05 and all tests were performed two-tailed. According to our hypotheses, we expected significant differences between the conditions regarding the ARCS dimensions as well as a significant correlation between the ARCS dimensions and learning success. These predictions were tested using t-tests for independent samples, using the dimensions attention, relevance, confidence and learning success as the dependent variable and condition (AR vs. CS) as the independent variable. As the data for satisfaction was skewed, the Kruskal-Wallis Test was used. No significant differences were found between the two conditions (see Figure 3).

As presented in Figure 3, the bivariate correlations between attention, relevance, satisfaction and learning outcome did not differ significantly from zero. However, the correlation between the learners' confidence and their learning outcome is significant, showing a moderate effect size that explains approximately 12% of the variance in the knowledge test. Since differences in the robot's behaviour did not influence attention, relevance, confidence or satisfaction, neither hypothesis 1 nor hypothesis 2 can be maintained. As the study cannot provide evidence for an influence of the robot's behaviour on the knowledge test, hypothesis 3 cannot be maintained either. The model was also tested using a path analysis, which led to the same results and is therefore not reported.

Additionally, we evaluated the robot's perceived intelligence and social presence, the enjoyableness and usefulness of the interaction with the learning environment, and whether the robot was suitable as learning companion. Table 3 shows the descriptive statistics for the users' perceptions. Ratings by participants are similar for both conditions. Mann-Whitney U tests did not show any significant differences (perceived intelligence, $U = 164.5$, $p = .48$, social presence, $U = 163.0$, $p = .46$, enjoyableness of interaction, $U = 153.0$, $p = .30$, usefulness of interaction, $U = 175.5$, $p = .69$, ability as learning companion, $U = 168.0$, $p = .55$). In general, the interaction with the robot was rated rather high in both conditions, with means between $M = 3.42$ and $M = 4.66$ on scales ranging from 1 to 6. On average, participants also stated that they felt "hardly" distracted by the robot.

6. Discussion

Contrary to our expectations, there were no significant differences between the experimental conditions regarding participants' motivation or learning success. The values of the ratings on the ARCS scales show that the manipulation did not lead to the intended results, although tending in the predicted direction. For example, the means for *attention* and *relevance* are slightly higher in the AR condition than those in the CS condition - vice versa participants in the CS condition rated their own *confidence* and *satisfaction* higher than those in the AR condition (see Table 2). It is concluded that the differences between the conditions were not strong enough to elicit bigger effect sizes.

In general, participants in both conditions felt relatively motivated, as can be seen from the descriptive values in the motivational scales (see Table 2). This is consistent with the findings of Chang et al. (2010), whereby robots in an educational context are perceived as motivating. However, it must be kept in mind that the novelty effect can play a role when new instructional methods are introduced (Kerres, 2001). This can lead to an increase in learners' motivation. To be able to eliminate the influence of this effect, further long-term studies are necessary.

In addition, the learning success for both groups was satisfactory: On average, participants solved nearly 65% of the exercises given in the language test (see Table 2). The relatively large number of correct answers leads to the impression that the learning goals have been reached.

Another noteworthy aspect is that, descriptively, participants in the CS condition achieved higher scores in the language test than those in the AR condition. A possible explanation is that the robot spent more time on individualized feedback in the CS condition. For example, the robot did not only tell the learners "That's correct" or "Wrong", but addressed the learner individually (e.g. "Take a closer look") depending on whether they gave correct or incorrect answers. Therefore, participants might have engaged with the correct answers in a more intensive way and therefore memorized them better. On top of that, negative feedback possibly led to an increase in the learners' ambition to solve subsequent exercises. This is supported by our observation that learning speed in the CS condition usually decreased following a negative comment made by the robot.

Comments given by the participants during or after the experiment also indicate that robot behaviour after the exercises (CS condition) bears great potential: In both conditions, participants stated that, instead of just being told about the incorrectness of the given answer, they would have liked the robot to *"[...] explain why the wrong answer isn't right"* in order to gain content-related feedback. More detailed feedback of this kind could possibly prevent similar mistakes in the next language test: *"An explanation why something is right or wrong would be helpful to understand the material and avoid making the same mistake again"*, as one of the participants stated. Thus the marginal difference in learning outcome between the two conditions could possibly have been increased by more detailed and personalized feedback in the CS condition.

Further, the study revealed a significant correlation between the learners' *confidence* and their learning success. This is consistent with Bandura's self-efficacy concept (Bandura, 1976), which states that the learners' subjective appraisal of their own competences is reflected in their later performance. Since the manipulation did not work and the robot's behaviour did not lead to increased levels of *confidence* in the CS condition compared to the AR condition, learners could not take advantage of this correlation in terms of differences in learning success. Nevertheless, the results of the experiment imply that successful manipulation would lead to increased learning success in the CS condition.

The knowledge test, which has been developed in cooperation with the teachers of the Spanish language courses, worked well. While a small number of items were slightly too easy or too difficult, the average item difficulty in the language test is .64 and therefore satisfying (Eid, Gollwitzer & Schmitt, 2010), and Cronbach's alpha indicates good reliability for the test. On top of that, participants stated that *"[...] the learning platform had appropriate and varying questions"* or *"meaningful exercises"*.

As stated in the results section, there were no significant differences in the rating of the robot between the conditions (see Table 3). Nevertheless, it can be noted that the robot scored higher than average on all of the scales. The rating concerning the quality of the interaction with the learning environment shows that the integration of the robot was meaningful. Participants noted that *"The study was fun"*; *"This is a great experiment and at the same time an opportunity to learn Spanish"*. Some participants considered the verbal feedback given by the robot as especially positive. This is in line with the modality effect (Mayer, 2010), which predicts that auditory information can be processed better than written information in cases where learners look at a picture, or written exercise, at the same time.

7. Conclusion and future work

This article investigates the effects of a social robot's motivational behaviour prior to (AR) vs. after (CS) several exercises in learning Spanish based on the ARCS model. A learning environment has accordingly been implemented that integrates a social robot with a visual interface that displays learning tasks.

The results do not show significant differences between the conditions, in terms of participants' motivation and learning success. The manipulation of the four motivational dimensions *attention, relevance, confidence* and *satisfaction* did not work as intended, although the groups ranked slightly higher on their respective scales. Furthermore, above-average means on the scales measuring motivation and learning success suggest that both conditions had a positive effect. Learning success in the CS group tended to be higher, which implies that motivational feedback after an exercise offers great potential. Therefore, we think our results encourage the view that motivational robot behaviour can improve learning outcomes in an educational context. Follow-up studies could clarify how to achieve the maximum benefit for learners using our learning environment.

We plan to extend our learning environment using the following aspects: the robot should refer to each task more individually and, particularly if the given answer was wrong, explain the mistake made by the user. Utilizing informative feedback of this kind, learners would have the chance to analyse the subject in more depth and learn from their mistakes.

Moreover, a comparison between a control group that interacts with the learning environment without a motivational robot, and a condition in which the robot motivates the learners prior to as well as after the exercise would allow evaluation of the general effect of the robot on motivation and learning outcome. Furthermore, a long-term investigation of this study would be highly beneficial, where learners would interact with the robot during multiple, time-intensive sessions. This could eliminate the assumption that positive learning effects have been triggered solely by the novelty effect (Kerres, 2001).

References

Admoni, H., & Scassellati, B. (2014). Role of Robots in Socially Assistive Applications. *International Journal of Social Robotics*.

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction, 16*, 389–400.

Bandura, A. (1976). Self-reinforcement: Theoretical and methodological considerations. *Behaviorism*, *4*(2), 135-155.

Bartneck, C., Kulic, D., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71-81.

Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior, 25,* 450-457.

Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education, 15*(2), 95-115.

Bernaus, M. (1995). *The role of motivating in the learning of English as a foreign language.* Bells: Barcelona English Language and Literature Studies.

Breazeal, C. (1998). A motivational system for regulating human-robot interaction. In *Aaai/iaai* (pp. 54-61).

Breazeal, C. (2003). Toward Sociable Robots. *Robotics and Autonomous Systems, 42*(3-4), 167-175.

Brown, L., Kerwin, R., & Howard, A. M. (2013). Applying behavioral strategies for student engagement using a robotic educational agent. In *Systems, Man, and Cybernetics (SMC), 2013 ieee International Conference on* (pp. 4360-4365). IEEE.

Carroll, J. B. (1973). Ein Modell schulischen Lernens. In W. Edelstein, & D. Hopf (Eds.), *Bedingungen des Bildungsprozesses* (pp. 234-250). Klett.

Chang, C. W., Lee, J. H., Wang, P. Y., & Chen, G. D. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Educational Technology and Society*, 13-24.

Cook, D. A., Beckman, T. J., Thomas, K. G., & Thompson, W. G. (2009). Measuring motivational characteristics of courses: Applying Keller's instructional materials motivation survey to a web-based course. *Academic Medicine, 84*, 1505-1509. doi:10.1097/ACM.0b013e3181baf56d

Deimann, M. (2002). Motivationale Bedingungen beim Lernen mit Neuen Medien. *Medienunterstütztes Lernen - Beiträge von der WissPro Wintertagung,* 61-70.

Eid, M., Gollwitzer, M., & Schmitt, M. (2010). *Statistik und Forschungsmethode*n. Weinheim: Beltz.

Fasola, J., & Mataric, M. (2012). Using Socially Assistive Human-Robot Interaction to Motivate Physical Exercise for Older Adults. *Proceedings of the IEEE*, *100*(8), 2512-2526.

Gara, N. (2001). Lifelong learning and technical and further education. In: D. Aspin, J. Chapman, M. Hatton, & Y. Sawano (Eds.), *International Handbook of Lifelong Learning*. *Springer International Handbooks of Education, Vol 6* (pp. 569-589). Springer, Dordrecht.

Gebhard, P., Mehlmann, G., & Kipp, M. (2011). Visual SceneMaker - A tool for authoring interactive virtual characters. *Journal on Multimodal User Interfaces*, *6*, 3-11.

Huang, B., & Hew, K. F. T. (2016). Measuring learners' motivation level in massive open online courses. *International Journal of Information and Education Technology*, *6*(10), 759-764.

Jung, Y., & Lee, K. M. (2004). Effects of physical embodiment on social presence of social robots. *Proceedings of PRESENCE*, 80-87.

Kanda, T., & Ishiguro, H. (2005). Communication robots for elementary schools. *Proceedings AISB05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, 54-63.

Kim, C., & Keller, J. M. (2008). Effects of motivational and volitional email messages (MVEM) with personal messages on undergraduate students' motivation, study habits and achievement. *British Journal of Educational Technology, 39,* 36-51. doi:10.1111/j.1467-8535.2007.00701.x

Keller, J., & Suzuki, K. (2004). Learner motivation and e-learning design: A multinationally validated process. *Journal of Educational Media*, *29*(3), 229-239.

Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Eds.), *Instructional Design Theories and Models - an Overview of their Current Studies* (pp. 383-434). Hillsdale, New Jersey: Erlbaum.

Keller, J. M., & Kopp, T. W. (1987). Motivational design of instruction. In C. M. Reigeluth (Eds.), *Instructional Design Theories and Models - an Overview of their Current Studies* (pp. 289-320). Hillsdale, New Jersey: Erlbaum.

Kennedy, J., Baxter, P., & and Belpaeme, T. (2015). The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning. In P*roceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 67–74). ACM.

Kerres, M. (2001). *Multimediale und telemediale Lernumgebungen. Konzeption und Entwicklung.* Oldenbourg, Muenchen: Walter de Gruyter.

Kidd, C. D., & Breazeal, C. (2008). Robots at home: Understanding long-term human-robot interaction. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3230–3235). IEEE.

Kim, C. ,& Keller, J. M. (2008). Effects of motivational and volitional email messages (MVEM) with personal messages on undergraduate students' motivation, study habits and achievement. *British Journal of Educational Technology*, *39*(1), 36-51.

Leyzberg, D., Spaulding, S. Toneva, M., & Scassellati, B. (2012). The physical presence of a robot tutor increases cognitive learning gains. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society (CogSci 2012)* (pp. 1882-1887).

Mandl, H., & Krause, U.-M. (2001). *Lernkompetenz für die Wissensgesellschaft*. (Forschungsbericht Nr. 145). München: Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie.

Mayer, R. (2010). *Multimedia Learning* (2nd ed.). Cambridge: Cambridge University Press.

Mori, J., Prendinger, H., & Ishizuka, M. (2003). Evaluation of an embodied conversational agent with affective behavior. In *Proceedings of the AAMAS03 Workshop on Embodied Conversational Characters as Individuals, Melbourne, Australia.*

Moriguchi, Y., Kanda, T., Ishiguro, H., Shimada, Y., & Itakura, S. (2011). Can young children learn words from a robot? *Interaction Studies, 12*(1), 107–118.

Mubin, O., Stevens, C. J., Shahid, S., Mahmud, A. A., & Dong, J. J. (2013). A review of the applicability of robots in education. *Journal of Technology in Education and Learning*, *1*, 1-7.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429–434.

Rheinberg, F. (2004). *Motivation.* Stuttgart: Kohlhammer.

Robins, B., Dautenhahn, K., Te Boekhorst, R., & Billard, A. (2005). Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? *Universal Access in the Information Society, 4*(2), 105–120.

Roesler, F. (2011). *Psychophysiologie der Kognition.* Heidelberg: Spektrum Akademischer Verlag.

Saerbeck, M., Schut, T., Bartneck, C., & Janse, M. D. (2010). Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1613-1622). ACM.

Schodde, T., Bergmann, K., & Kopp, S. (2017). Adaptive robot language tutoring based on bayesian knowledge tracing and predictive decision-making. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 128–136). ACM.

Small, R. V. (1999). Motivation in instructional design. *Educational Media and Technology Yearbook*, 89-92.

Song, S. H., & Keller, J. M. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. *Educational Technology Research and Development*, *49*(2), 5-22.

Su, C. H., & Cheng, C. H. (2015). A mobile gamification learning system for improving the learning motivation and achievements. *Journal of Computer Assisted Learning, 31*, 268-286. doi:10.1111/jcal.12088

Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-*

Robot Interaction, 1(1), 78-95.

Ushida, H. (2010). Effect of social robots behavior in collaborative learning. *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction,* 195-196. IEEE.

Visser, J., & Keller, M. (1990). The clinical use of motivational messages: an inquiry into the validity of the ARCS model of motivational design. *Instructional Science, 19*, 467-500.

Yang, S. C., & Chen, Y. J. (2007). Technology-enhanced language learning: A case study. *Computers in Human Behavior*, *23*, 860-879.

Zimmerman, B. J (2011). Motivational Sources and Outcomes of Self-Regulated Learning and Performance. In D. H. Schunk and B. J. Zimmerman (Eds.), *Handbook of Self-Regulation of Learning and Performance,* 49-64.

Tables

Table 1

Reliabilities and Number of Items per Scale

	Cronbach's Alpha	Number of items
Attention	.86	10
Relevance	.76	9
Confidence	.72	8
Satisfaction	.88	7
Enjoyableness of interaction	.88	7
Usefulness of interaction	.94	4
Perceived intelligence	.80	5
Social presence	.79	6
Ability as learning companion	.93	5
Knowledge test	.79	23

Table 2

Means and Standard Deviations of the ARCS Dimensions and the Knowledge Test

	Attention		Relevance		Confidence		Satisfaction		Knowledge test	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
AR (*n* = 20)	4.20	.83	4.58	.55	4.37	.70	4.61	.88	62.81	17.08
CS (*n* = 19)	3.92	.98	4.40	.78	4.40	.72	4.64	.90	66.78	12.90

Note. Values range from 1 to 6, except for the knowledge test's values ranging from 1 to 100. AR refers to the attention and relevance condition, CS to the confidence and satisfaction condition.

Table 3

Means and Standard Deviations of Perception of the Robot and the Interaction with the Learning Environment

	Perceived Intelligence		Social Presence		Enjoyableness of interaction		Usefulness of interaction		Learning Companion	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
AR (n=20)	4.48	.73	46.45	23.96	4.66	.91	3.53	1.40	4.00	1.35
CS (n=19)	4.62	.60	42.16	18.95	4.35	1.02	3.43	1.46	3.67	1.53

Note. Values range from 1 to 6, except for social presence's values ranging from 1 to 100. AR refers to the attention and relevance condition, CS to the confidence and satisfaction condition.

Figure(s)

Scaffolding of Motivation in Learning using a Social Robot – Figures 1

Figures

Figure 1. Technical setup of the learning environment.

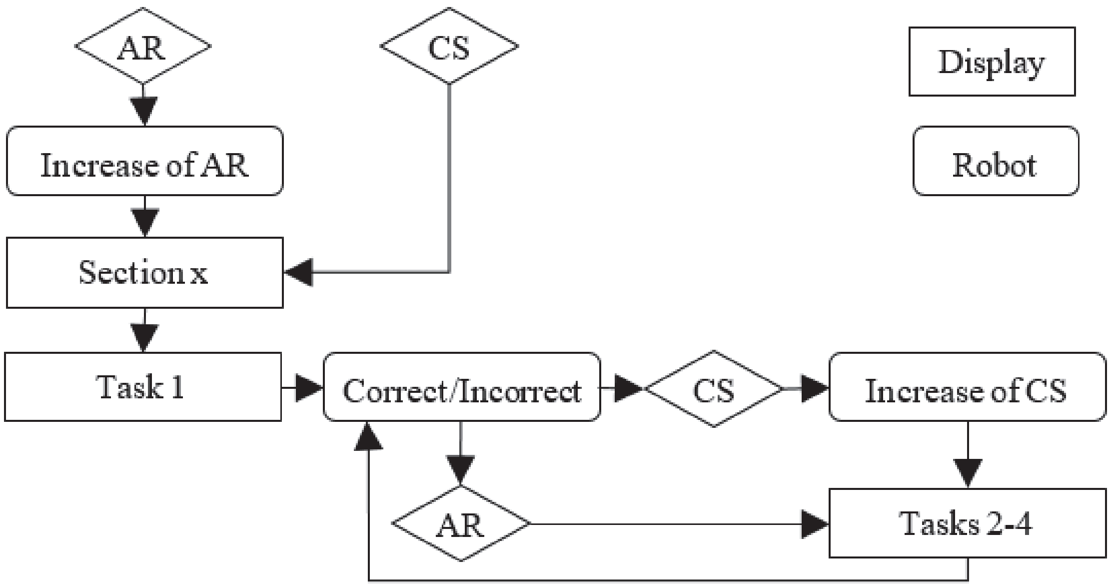

Figure 2. Procedure of interaction with the learning environment for one session (6 sessions in total). Squared boxes (Display) refer to the leaning material shown on a screen. Round boxes (Robot) refer to the motivational behaviour of the robot.

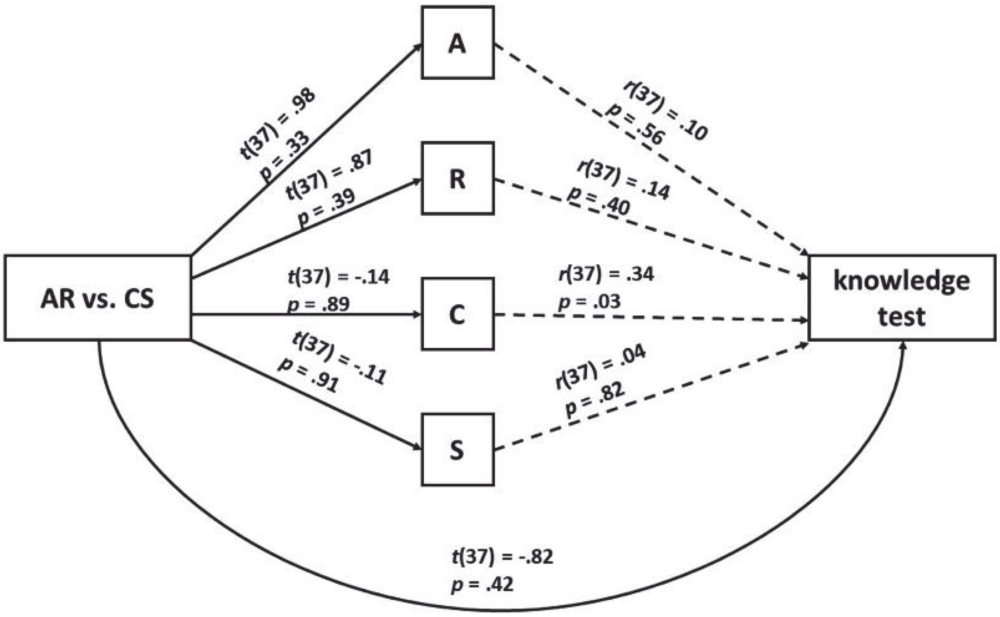

Figure 3. Results of the statistical tests. Connected lines show the comparison of means, dashed lines indicate Pearson's correlation.

Acknowledgements: We would like to thank the heads of the Spanish language department at the language Centre at Würzburg University, Ángel Curbelo and Mercedes Rodríguez Castrillión, for their help and cooperation.