Using Self-Determination Theory in Social Robots to Increase Motivation in L2 Word Learning

Peggy van Minkelen
Tilburg University
Tilburg, the Netherlands
p.vanminkelen@uvt.nl

Carmen Gruson
Tilburg University
Tilburg, the Netherlands

Pleun van Hees
Tilburg University
Tilburg, the Netherlands

Mirle Willems
Tilburg University
Tilburg, the Netherlands

Jan de Wit
Tilburg University
Tilburg, the Netherlands
j.m.s.dewit@uvt.nl

Rian Aarts
Tilburg University
Tilburg, the Netherlands
A.M.L.Aarts@uvt.nl

Jaap Denissen
Tilburg University
Tilburg, the Netherlands
jjadenissen@gmail.com

Paul Vogt
Tilburg University
Tilburg, the Netherlands
p.a.vogt@uvt.nl

ABSTRACT

This study presents a second language word learning experiment using a social robot with motivational strategies. These strategies were implemented in a social robot tutor to stimulate preschool children's intrinsic motivation. Subsequently, we investigated their effect on children's task engagement and word learning performance. The strategies were derived from the Self-Determination Theory, a well-known psychological theory that assumes that intrinsic motivation is strongly related to the fulfilment of three basic human needs, namely the need for autonomy, competence, and relatedness. We found an increase in the strength and duration of task engagement when all three psychological needs were supported by the robot. However, no significant results for learning gains were observed. Our intervention appears a promising method for improving child-robot interactions in educational settings, especially to sustain in long-term interactions.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; • **Applied computing → Interactive learning environments**.

KEYWORDS

motivation, self-determination theory, task engagement, second language learning, human-robot interaction, robot tutor

ACM Reference Format:
Peggy van Minkelen, Carmen Gruson, Pleun van Hees, Mirle Willems, Jan de Wit, Rian Aarts, Jaap Denissen, and Paul Vogt. 2020. Using Self-Determination Theory in Social Robots to Increase Motivation in L2 Word Learning. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20), March 23–26, 2020, Cambridge, United Kingdom.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3319502.3374828

HRI '20, March 23–26, 2020, Cambridge, United Kingdom

ACM ISBN 978-1-4503-6746-2/20/03...$15.00
https://doi.org/10.1145/3319502.3374828

1 INTRODUCTION

In recent years there has been an increasing interest in developing social robots for education [3, 6, 15], and more specifically for children's second language learning (henceforth L2) [4, 10, 14, 16, 26, 29]. There are various reasons for this increased interest. For example, robots have the advantage to provide one-on-one tutoring to improve the learning outcome by assisting a human teacher, and tailoring the learning experience to the child [3, 15, 27]. This way, a robot can remain inside the child's Zone of Proximal Development [30]. However, experimental studies in L2 learning using social robots have shown mixed results regarding learning outcomes. For example, Westlund et al. [31] and Van den Berghe et al. [25] found no differences in L2 learning outcome between children in an L2 learning task with human peers compared to robot peers. Contrary, Alemi, Meghdari, and Ghazisaedy [1] found an improvement in learning outcome when a robot was present as a teaching assistant compared to when the robot was not present.

In order for social robots in education to be successful, long-term interactions are essential, which means that children need to remain motivated to engage in the learning task provided by the robot over multiple sessions. Motivation in children's learning from robots has been found to correlate with learning outcome [2], enhanced interest and concentration towards the learning task compared to other educational forms [11], increased confidence in children [32], and reduced anxiety in children's L2 learning when the robot is framed as a peer tutor [2]. From psychological studies it is known that motivation –and specifically intrinsic motivation– assists children's engagement in learning [21, 23]. Blumenfeld, Kempler, and Krajcik, in their study within the field of educational science [5], demonstrated that task engagement mediates the effect of motivation on learning outcome. They suggested a positive feedback loop, starting when motivation increases interest, which then in turn enlarges task engagement and subsequently increases children's skills. These increased skills could then lead to a further increase in motivation and sustained task engagement. In other words, since motivation is a prerequisite for task engagement [23], and because engagement increases learning performance [5], it is imperative to motivate children during their learning task.

That said, various studies have found that task engagement wears off over time, either within a single session with a robot [7] or between multiple sessions [14, 22]. This decrease in children's task engagement with robots has been attributed to the novelty effect [19]. The novelty effect entails that children are initially excited to interact with a robot, resulting in high task engagement. But when children become familiarised with the robot, the novelty effect wears off, and children's task engagement tends to decrease.

While various studies have investigated the effect of motivation on L2 learning with social robots [2, 11, 15, 16, 32], few studies have designed strategies for increasing learner's motivation explicitly based on fundamental psychological theories about motivation (e.g. [16]). Most studies incorporate one or more strategies that might be correlated with intrinsic motivation without a specific aim to investigate this possible effect. In an important exception, Kennedy and colleagues [16] have investigated the effect of verbal immediacy on children's learning performance. They have based their model on a theory about 'psychological availability' as a way to enhance motivation [33], and implemented this by having the robot relate to children using verbal utterances that personalise the interaction, e.g. by using the child's name, revealing personal information and using more inclusive vocabulary such as 'we', 'our', and higher praise in feedback. While children perceived the robot that used higher availability to score higher on verbal immediacy, these children did not learn better than the children who learned from a robot with low availability [16]. So, it remains unclear how a social robot can structurally motivate children for longer periods of time, such that the learning performance improves.

One influential theory that addresses children's motivation and engagement in learning is the Self-Determination Theory (henceforth SDT) [8, 21]. SDT contains two sub theories; the Cognitive Evaluation Theory (CET) and the Basic Psychological Need theory (BPNT). CET states that intrinsic motivation thrives mainly on three intertwined psychological needs. First, autonomy comprises performing a task based on one's own volition. Second, competence comprises the self-perceived ability to learn new things and receive feedback, which will only enhance intrinsic motivation once accompanied by a sense of autonomy. Third, BPNT argues that relatedness, i.e. a sense of belongingness, plays a substantial role in maintaining intrinsic motivation.

In an educational setting, autonomy can be satisfied by children's need to experience out of their own interest, providing children with choice and minimizing external pressure [8, 21, 28]. The psychological need for competence within an educational setting, can be satisfied by providing children with an appropriately challenging learning task, and with appropriate and/or positive feedback that will affect their feeling of being capable to fulfil the task [8, 21]. Finally, the need for relatedness can be satisfied by appreciating the child, ensuring that the child feels connected to others, in order for the child to internalize the learning material [8, 21, 24, 28]. To sum it up, in order for children to increase intrinsic motivation, and consequently task engagement in a learning environment, their needs for autonomy, competence, and relatedness need to be met. Hence, social robot tutors in an L2 word learning task might become more effective when supporting these needs.

De Wit et al. [7] is one example in which some motivational strategies were investigated, although without explicitly referring to motivation as an objective. Instead, they investigated the effect of using gestures and adaptive tutoring on children's L2 learning performance and engagement. In this study, preschool children played the game *I spy with my little eye* in a single session using a NAO robot as a tutor, in a 2 (adaptive tutoring vs. random tutoring) by 2 (gestures vs no gestures) between-subjects design. They implemented an experimental condition, which they called an "adaptive strategy", in which the robot adapted the tasks based on its knowledge of the child's performance on previous attempts. They compared this condition to the tutoring strategy in which the robot randomly presented the next learning task. They found that children's engagement decreased over time in all conditions, but the decrease in engagement was less steep when the robot used the adaptive tutoring strategy. One could argue that adaptive tutoring is a strategy that fulfils a child's need for competence, because it reacts to a child's learning performance, albeit implicitly, and attempts to present an optimal level of challenge in the tasks. So it seemed that adaptive tutoring would motivate children to maintain their task engagement for a prolonged period of time. However, like the verbal immediacy in Kennedy et al.'s study [16], the adaptive tutoring strategy did not affect the learning outcome. Yet the use of gestures did improve both engagement and learning gain [7]. Hence, it remains unclear how motivational strategies to support the psychological needs for autonomy, competence, and relatedness, can be best implemented in a robot.

This paper reports new insights into whether a social robot can satisfy these psychological needs for autonomy, competence, and relatedness, based on the SDT theory [8, 21]. The current study investigates whether a social robot can satisfy these psychological needs and whether these motivational capacities can maintain children's task engagement in an L2 word learning task over time, and, as a result, increase L2 word learning. To this aim, we carried out an experiment that takes the baseline condition (random tutoring without using gestures) from De Wit et al. [7] as a control condition and compares this model with two conditions that apply motivational strategies aimed to implement the SDT theory. One condition implements SDT partially, only applying strategies to satisfy autonomy and competence; the other also applies strategies to satisfy relatedness. Based on the findings in the aforementioned studies, we hypothesized that:

- H1 Children's engagement in an L2 word learning task is stronger when the needs for autonomy, competence and relatedness are, at least partially, satisfied, compared to the control condition (H1a) and children's task engagement is stronger when all three needs are satisfied, compared to the condition where only the needs for autonomy and competence are satisfied (H1b).
- H2 Children remain more engaged over time towards an L2 word learning task when the needs for autonomy, competence and relatedness are, at least partially, satisfied, compared to the control condition (H2a) and they remain more engaged when all three needs are satisfied, compared to the condition where only the needs for autonomy and competence are satisfied (H2b).
- H3 Children will learn more target words in an L2 word learning task when the needs for autonomy, competence and relatedness are, at least partially, satisfied, compared to the control condition (H3a) and children will learn more target words when all three needs are satisfied, compared to the condition

where only the needs for autonomy and competence are satisfied (H3b).

H4 Children will remember the target words in an L2 word learning task better when the needs for autonomy, competence and relatedness are, at least partially, satisfied, compared to the control condition (H4a) and children will remember the target words better when all three needs are satisfied, compared to the condition where only the needs for autonomy and competence are satisfied (H4b).

H5 Children's task engagement during an L2 word learning task positively correlates with their learning gain.

2 LESSON DESIGN

In an experiment with Dutch preschool children during an L2 word learning task, we incorporated motivational support interventions based on CET and BPNT in a social robot. Our choices for the support interventions were based on the description of the three psychological needs. We assessed the applicability of these support inventions in a short L2 word learning task with a social robot. Other than these added support interventions, our design stayed close to De Wit et al. [7], since we incorporated their data of the "random tutoring without gestures" condition as our control condition. We used this condition, because both their "adaptive tutoring" and their "iconic gesture" conditions had effects on children's task engagement, and thus seemed to contribute to motivational capacities. In our experimental condition, we implemented motivational strategies to fulfil the needs for autonomy, competence and relatedness. We assessed children's task engagement towards the L2 learning task as well as the effect of motivational support interventions on their learning outcome.

2.1 Target Words

Since we used the data of study by De Wit et al. [7] as our control group, we adopted the six target words from their study. During the experiment, the robot tried to teach Dutch children (5 to 6 years old) six English words. These words were identical as in De Wit et al. [7]: bird, monkey, horse, ladybug, hippo, and chicken. Each word was associated with a drawing of the animal displayed on a tablet during the word learning sessions (cf. Figure 1).

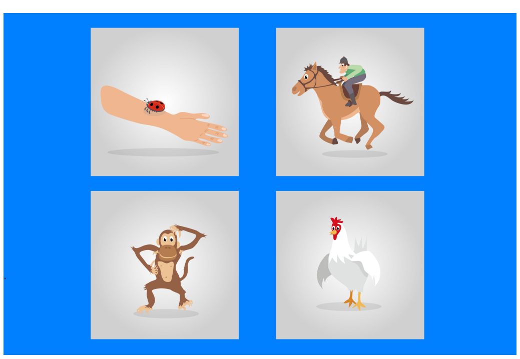

Figure 1: Target word and three distractors presented on the tablet during the training phase.

2.2 Lesson

We adopted de Wit et al.'s condition without gestures and without adaptive tutoring, because this condition appeared least motivating for children [7]. We purposely selected this monotonous task to investigate the possible effects of the incorporated motivational support on children's motivation to perform 30 rounds of the same task. The experiment consisted of 30 rounds of the game *I spy with my little eye*, in which the robot randomly chose one of the six target words and then said *Ik zie, ik zie wat jij niet ziet en het is een* [L2 target word] (in English: *I spy with my little eye a* [L2 target word]). Only the L2 target word was spoken in English, the rest was spoken in Dutch. Pictures of the target word and two to three distractors were presented to the child on the tablet (see Figure 1). The child had to select a picture on the tablet, which was then highlighted. If the child selected the target, a green, happy smiley appeared on the tablet with verbal feedback in the first language (henceforth L1), Dutch (e.g. "Goed gedaan, het Engelse woord voor [L1 target word] is [L2 target word]"; In English: "Well done, the English word for [L1 target word] is [L2 target word]"). However, if the child selected the wrong picture, a red, sad smiley appeared on the tablet with verbal feedback in Dutch (e.g. "Dat was een [selected L1 target word], maar ik zag een [L2 target word]. [L2 target word] is het Engelse woord voor [L1 target word]". In English: "That was a [selected L1 word], but I saw a [L2 target word]. [L2 target word] is the English word for [L1 target word]"). Next, the target word was presented again, only this time with just one distractor.

2.3 Motivational Strategies

The need for autonomy was implemented by granting choice to the child every fifth round of the game in the experiment. The child could choose between playing a game showing three pictures (one target word and two distractors) or a game showing four pictures (one target word and three distractors). The choice, however, did not affect the learning task, merely the number of distractors presented; an aspect that the child was not told of. In all three conditions, the need for autonomy was also supported by presenting a green smiley which the child could press to start the lesson, and a red smiley which the child could press to indicate he/she had not understood the instruction.

The need for competence was implemented by providing one of six positive feedback sentences when the child chose the right picture. Five of these sentences were identical to the study of De Wit et al. [7], except the sentence "Wow, well done. That is right". Besides positive feedback sentences, the robot added supportive sentences when providing corrections after the child chose the wrong picture (e.g. "Dat was moeilijk. Probeer het nog eens."; In English: "That one was difficult. Let's try again."). The study by De Wit et al. [7] only corrected the answer, using the same sentence every time and did not incorporate supportive sentences. Additionally, performance-contingent rewards would appear after the child selected a picture, i.e. a green, happy smiley appeared if the child selected the target (see Figure 2, left), whereas a red, sad smiley appeared if the child selected the wrong picture (see Figure 2, right).

The need for relatedness was implemented firstly by placing the robot close to the child at a 45-degree angle to create a feeling of being peers, as opposed to the robot facing the child at a 180-degree

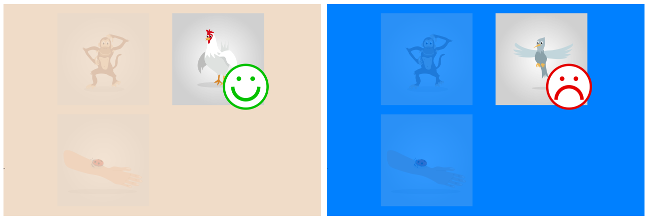

Figure 2: Performance-contingent rewards after the child selected the target (left) or the wrong picture (right). Left shows the *AC-condition* with the default colour of the tablet, right shows the *ACR-condition* where the child chose 'blue' as favourite colour.

angle in the conditions without relatedness. Secondly, the robot asked the child what his/her favourite colour was, as ice breaking behaviour, after which the background colour of the tablet changed into the child's favourite colour. Thirdly, the robot regularly called the child by name when providing positive feedback once every three rounds of feedback (e.g. "Well done [name child]"), and in every instance of supportive feedback. Finally, the robot suggested the child to fist bump, a gesture with a similar meaning as a handshake or high-five, which children use among peers. The fist bump provides a positive and supportive interaction with the robot, and was found to positively affect children's motivation [20]. The fist bump was executed every sixth round (see Figure 3).

Figure 3: The robot suggests a fist bump as measure to implement relatedness.

While the control condition from De Wit et al. [7] did not include specific motivational supportive interventions, some motivational aspects did occur in the original study. For example, the robot addressed each child by his or her name, but only once at the start of the experiment. Also, the robot provided feedback at the end of each game, but always in the same way and not as varied and elaborate as in the current study.

3 METHOD

An experiment was conducted in a between-subjects design to investigate the effect of motivational capacities of a social robot on children's task engagement and learning gain. The study consisted of three conditions:

(1) *AC-condition.* Motivational supportive interventions based on CET, in which the two needs for autonomy and competence are supported by the robot ($N = 23$).
(2) *ACR-condition.* Motivational supportive interventions based on BPNT, in which the needs for autonomy, competence, and relatedness are supported by the robot ($N = 26$).
(3) *Control condition.* No specific motivational supportive interventions are supported by the robot. This condition is based on the "random tutoring without gestures condition" of De Wit et al. [7] ($N = 16$).

3.1 Participants

The experiment was conducted at three primary schools in Tilburg, the Netherlands. A total of 49 native Dutch speaking children of five and six years old were recruited by contacting the schools after which the school and parents/ caretakers were informed by letter about the experiment and were sent consent forms. The study received ethical approval from the Research Ethics Committee of Tilburg University, as part of the L2TOR project. We excluded two children who scored less than five words correct in the Dutch pretest and one child who scored more than five words correct in the English pretest. Additionally, we excluded one child who was not a native speaker of Dutch, and one child who did not complete the whole experiment. The remaining 49 children had a mean age of 5;8 years old (69.04 months; $SD = 5.13$), of which 23 children were female and 26 children were male. The 49 children were evenly distributed among the *AC-condition* (Mean age 68.91 months; $SD = 5.23$; 12 female; 11 male) and the *ACR-condition* (Mean age 69.15 months; $SD = 5.14$; 11 female; 15 male). In the *Control condition* a total of 16 children had a mean age of 5.04 years old (60.44 months; $SD = 7.26$; 8 female; 8 male). All 65 children within the study had an average age of 5.6 years old (66.9 months; $SD = 6.7$; 31 female; 34 male).

3.2 Materials

A Softbank Robotics NAO robot was used in the experiment, since previous research has shown that preschoolers feel comfortable interacting with NAO and perceive NAO as a peer [10, 12]. Moreover, NAO was similar to the robot used in the study by De Wit et al. [7], which meant we could use their data as control condition.

The interaction between the child and the robot was recorded with two cameras: one facing the child to provide adequate footage to assess task engagement of the child with the robot, and one camera at a 90-degree angle to provide a more complete overview.

A laptop was used to display the images of the six target words during the pretest and post-test (see Figure 4). Pre-recorded fragments of the animal names of the six target words were pronounced first in Dutch (L1) and then in English (L2), by a bilingual female speaker. Additionally, the laptop was used in the concept-binding phase in which the images were displayed one by one, accompanied by the pre-recorded fragment, "Look, this is a [L2 target word]. Do you see the [L2 target word]? Click on the [L2 target word]". The images used in the pretest and post-test, the concept-binding, and the training were obtained from the study of De Wit et al. [7]. The images that were presented during the pretest, concept-binding, and post-test were different from the images used in the training

Figure 4: Six target words as presented on the laptop during the pretest and post-test.

phase (see Figure 1), to make sure the child obtained the concepts of L2 words instead of mapping L2 words to images.

A tablet (Microsoft Surface Pro) was placed on the table in a slightly tilted position. We used the tablet to display the learning task on the tablet and record the child's responses, since automatic speech recognition of NAO is not reliable [17], and the robot cannot monitor the child's movements through space autonomously [4].

3.3 Procedure

The experiment was conducted in a separate classroom dedicated to the experimental procedure. Two experimenters were present at all times. One experimenter was responsible for the child, i.e. getting the child out of the classroom, approaching NAO together and discussing its appearance, explaining the procedure of the experiment, and bringing the child back to the classroom. The other experimenter was responsible for the set-up of the experiment. Children were told that the experimenters were occupied to avoid them seeking feedback or help. The experiment consisted of six phases, i.e. a group introduction, a Dutch and English pretest, concept binding, experimental phase, an immediate post-test and a delayed post-test. The different phases will be explained below.

3.3.1 Group Introduction. The robot was introduced to the children during a group introduction by one of the experimenters, following Fridin [9], who found that children who first met the robot during a group introduction encountered a more positive interaction with the robot. The group introduction was similar to the group introduction of De Wit et al. [7], and consisted of a background story of the robot named 'Robin', to frame Robin as a peer. In sum, the story described that Robin was seven years old, that he was looking for new friends, and that he was learning English words in preparation for his upcoming holiday to England. Subsequently, the experimenter pointed out the similarities in physique between Robin and humans, such as having hands and eyes. Next, Robin performed three dances, after which the children could shake hands with Robin. Finally, the children put Robin to bed. The group introduction was scheduled a few days before the experiment.

3.3.2 Pretest. A Dutch (L1) pretest, followed by an English (L2) pretest was conducted to test the child's prior knowledge of the six target words. Both tests consisted of images of the six target words which were presented all at once (cf. Figure 4). For the L1 pretest, a pre-recorded voice pronounced the six target words in Dutch (L1) and the child was asked to click on the image which corresponded to the pronounced word. For the L2 pretest, the same was done, but with the target words pronounced in English.

3.3.3 Concept Binding. After the two pretests, the child proceeded with the concept binding phase, providing the child with a correct match between the L2 target word and the corresponding image. This phase was necessary since the pretests did not provide the child with feedback and therefore the child could assume being right about a certain concept linked to a word when in fact, he/she was wrong. While the laptop displayed one image of a target word, a pre-recorded voice pronounced; "Kijk, dit is een [L2 target word]. Zie jij de [L2 target word]? Tik maar op de [L2 target word]." In English: "Look, this is a [L2 target word]. Do you see the [L2 target word]? Click on the [L2 target word]".

3.3.4 Experimental Phase. The experimental phase consisted of two test rounds, one in Dutch (L1) and one in English (L2), followed by the actual experiment of 30 rounds of *I spy with my little eye*, as described in Section 2.2. The robot said *Ik zie, ik zie wat jij niet ziet en het is een* [L2 target word]; in English: *I spy with my little eye a* [L2 target word], after which the child had to select a picture. Each round, the target word was randomly chosen, but each of the six words was presented in five different rounds for an equal balance between target words. In all three conditions, children played the same 30 rounds, however, depending on the condition they were assigned to, different motivational strategies were implemented in the robot, as described in section 2.3. The number of times a target word was presented varied across children, because the robot presented a target word again if the child had not selected the target word correctly.

3.3.5 Immediate and Delayed Post-test. Immediately after the experimental phase, the child completed a post-test to measure the immediate L2 learning gain. After approximately one week, the child completed the delayed post-test, to measure long-term retention of the L2 target words. The immediate and delayed post-tests were identical to the English pretest.

3.4 Measures

To investigate whether motivational support interventions affect children's task engagement in the L2 word learning task over time, we measured task engagement of the child on a 5-point scale with 0.5 intervals (0 is not engaged at all, 5 is maximally engaged)[1]. To this aim, we coded 2-minute video clips, starting at the fourth round and the twenty-fourth round. We used the fourth round so the child could get familiarized with the robot, and we used the twenty-fourth round to ensure we could code at least two minutes, since the duration of the rounds varied across participants. Task engagement comprises of robot engagement by listening to and looking at the robot, since the robot is the instructor, as well as engagement towards the task by selecting images on the tablet. Children were coded as not engaged when, e.g. they were not paying attention to the robot or they were not executing the task. Contrary, children were coded as maximally engaged when, e.g. they continuously paid attention to the robot or they showed a

[1] https://github.com/l2tor/codingscheme

high focus on completing the task. We re-coded engagement of the control group using this same coding scheme, as the study of De Wit and colleagues [7] used a different method. Two raters were trained in coding engagement using this coding scheme. They both rated 25% of the video fragments and we calculated their inter-rater reliability through the intraclass correlation coefficient [18] based on a single rater, consistency, two-way random effects model. The 95% confidence interval was [.60, .89], which can be considered moderate–good [18].

Learning gain was measured by subtracting the results of the English pretest from the results of the post-test (immediate learning gain), and by subtracting the results from the English pretest from the results of the delayed post-test (delayed learning gain). All scores ranged from 0 (no learning gain) to 6 (maximum learning gain) for a child with zero pretest knowledge.

3.5 Analysis

To examine the effects of motivational support interventions on children's task engagement and learning gain, we carried out four one-way between-subjects ANOVAs with planned contrasts with the three conditions (*ACR-condition*, *AC-condition* and *Control-condition*) as independent variables. To measure the effect on the overall task engagement, we took the mean task engagement per child from the fourth and twenty-fourth round as dependent variable. To measure the effect on task engagement over time, we took the difference between the fourth and twenty-fourth round as dependent variable. To measure the effects on learning gain, we took the differences between the immediate post-test and pretest, and between the delayed post-test and pretest as dependent variables. For all analyses, Levene's test indicated equal variances between independent variables across the three conditions ($p > .05$). To examine a possible relationship across the three different conditions between task engagement and immediate learning gain on the one hand, and between task engagement and delayed learning gain on the other, we carried out a series of Pearson's correlation tests.

4 RESULTS

4.1 Task Engagement

Figure 5 shows that, overall, mean task engagement of children was higher in the *ACR-condition* ($M = 3.48, SD = 0.58$), compared to the *AC-condition* ($M = 3.03, SD = 0.58$) and the *Control-condition* ($M = 3.09, SD = 0.63$). The one-way ANOVA with planned contrasts revealed that there was a significant effect of condition on the mean task engagement, $F(2, 62) = 4.05, p = .022, \eta_p^2 = .12$. Since the mean task engagement in the control group was slightly negatively skewed (z-score skewness = -2.01) we performed a bootstrapped analysis. Planned contrast revealed that there was no significant difference when comparing the *AC-condition* and *ACR-condition* with the *Control-condition*, $t(62) = 0.96, p = .342$. However, children in the *ACR-condition* were significantly more engaged toward the robot than children in the *AC-condition*, $t(62) = 2.65, p = .010, r = 0.34$. So, the data did not support hypothesis H1a, though it supported hypothesis H1b.

Additionally, we investigated the effect of motivational support interventions of the robot on keeping children engaged towards the learning task. Figure 5 shows that children's engagement dropped

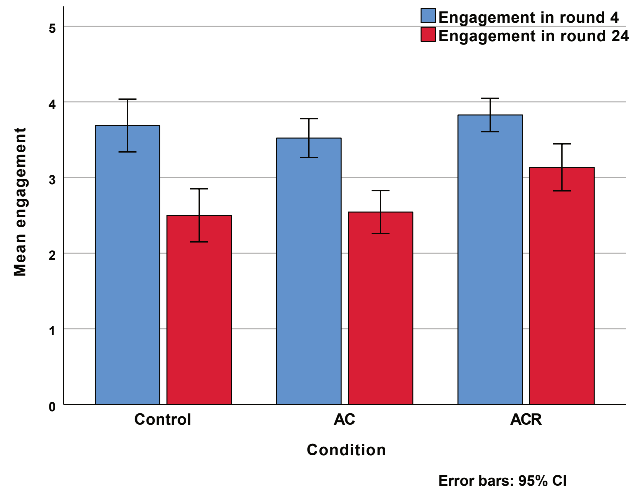

Figure 5: Mean task engagement of the fourth and twenty-fourth round between the three conditions.

between the fourth and the twenty-fourth round in all three conditions. However, the difference in engagement between the fourth and the twenty-fourth round was the smallest in the *ACR-condition*.

A dependent samples t-test revealed that on average all children ($N = 65$) showed significantly stronger task engagement in the fourth round ($M = 3.69, SD = 0.60$) than in the twenty-fourth round ($M = 2.77, SD = 0.76$), $Mdif = 0.92, t(64) = 12.95, p < .001$, 95% CI [0.77, 1.06], with a large Cohen's d of 1.35, as expected based on the literature. The figure clearly shows that task engagement dropped much less in the *ACR-condition* ($M = -0.69, SD = 0.66$), compared to the *AC-condition* ($M = -0.98, SD = 0.49$) and the *Control-condition* ($M = -1.19, SD = 0.36$). A second one-way ANOVA with the difference in task engagement between the fourth and the twenty-fourth round as dependent variable revealed a significant effect of the type of condition on remaining engaged, $F(2, 62) = 4.37, p = .017, \eta_p^2 = .12$. Planned contrast revealed that the *ACR-condition* and *AC-condition* task engagement dropped significantly less than in the *Control-condition*, $t(62) = 2.26, p = .028, r = 0.35$. There was, however, no significant difference between the *ACR-condition* and the *AC-condition*, $t(62) = 1.84, p = .070$. So, the data supported hypothesis H2a, suggesting that children in the *ACR-condition* and *AC-condition* remain more engaged over time towards an L2 word learning task compared to children in the *Control-condition*. However, the data did not support hypothesis H2b.

4.2 Learning Gain

We examined the effect of motivational support interventions of the robot on children's learning gain. Figure 6 shows that children's immediate and delayed learning gain of L2 words were the highest in the *ACR-condition*.

To test the effect of motivation on learning gain between the three conditions, a one-way between-subjects ANOVA was performed with the learning gain between the immediate post-test and the pretest as dependent variable. Although we found that the overall learning gain of children was higher in the *ACR-condition* (M = 1.35, SD = 1.26), compared to the *AC-condition* (M = 0.87, SD = 1.45) and the *Control-condition* (M = 0.94, SD = 1.44), these differences were not significant, $F(2, 62) = 0.84, p = .436$. So, the results did not support hypotheses H3a and H3b.

Similarly, when examining the learning gain between the delayed post-test and the pretest, we found that the overall learning gain of children was higher in the *ACR-condition* (M = 1.50, SD = 1.30), compared to the *AC-condition* (M = 1.04, SD = 1.58) and the *Control-condition* (M = 0.88, SD = 1.26), but again these differences were not significant, $F(2, 62) = 1.17, p = .317$. All in all, the data did not support hypothesis H4a and hypothesis H4b.

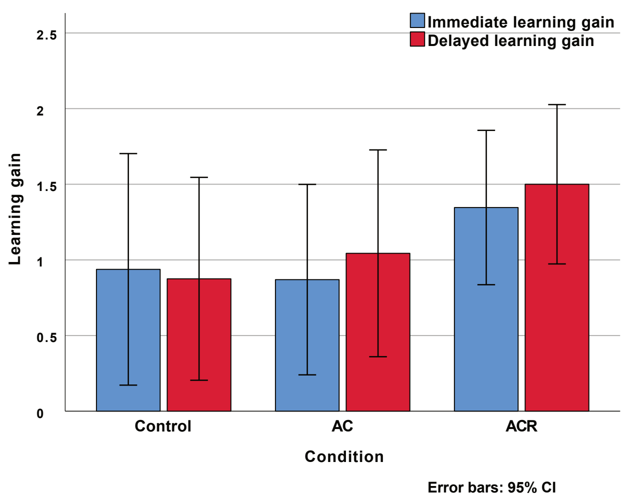

Figure 6: Average learning gains with 95% confidence intervals of the immediate post-test (blue bars) and delayed post-test (red bars) between the three conditions.

4.3 Correlation Task Engagement and Learning Gain

To test hypothesis H5, we analysed the correlation between the mean task engagement (M = 3.23; SD = 0.62) and immediate learning gain (M = 1.08; SD = 1.37). There was no significant correlation, Pearson $r_s = .05, p = .676$. When examining the correlation of mean task engagement and immediate learning gain for each of the three different conditions separately, we found a positive significant correlation only in the *AC-condition*, Pearson $r_s = .50, p = .017, r2 = .24$. Increases in engagement were correlated with increases in learning gain (see Figure 7).

Furthermore, we analysed if there was a correlation between mean task engagement (M = 3.23; SD = 0.62) and delayed learning gain (M = 1.18; SD = 1.40). There was no significant correlation, Pearson $r_s = .05, p = .692$, nor within one of the three conditions (all $ps > .05$). Exploring potential correlations further, we found no significant correlations between task engagement in the fourth or twenty-fourth round and immediate learning gain, or with longer-term learning gain. So, our correlation analyses did not support hypothesis H5.

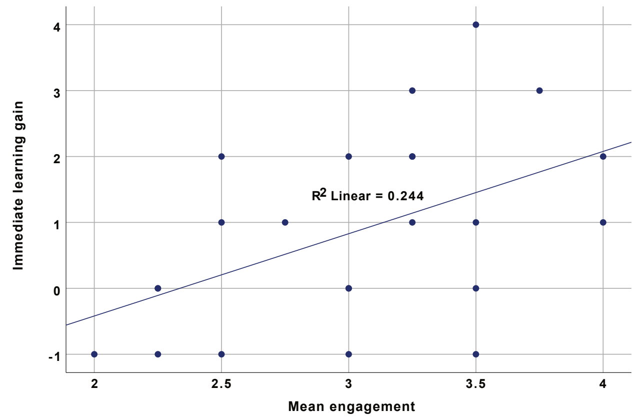

Figure 7: Scatterplot of immediate learning gain and mean engagement in the AC-condition.

5 DISCUSSION

Our study is a first attempt to implement an intervention to support the psychological needs for autonomy, competence, and relatedness based on the Self-Determination Theory [8, 21] by a social robot. The most important finding to emerge from our analysis is that motivational strategies which satisfy all three psychological needs, seem to keep children engaged over a longer period of time. The finding that a condition that supported all three psychological needs had a larger effect on engagement than a condition with only competence and autonomy implemented, could be attributed to the added value of satisfying the need for relatedness. It seems that children benefit from the sense of belongingness, which is in line with the Basic Psychological Need theory [8, 21]. The strategy to satisfy the need for relatedness was supported by personalizing the interaction (i.e. the robot addressing the child by name and changing the background colour of the tablet into the child's favourite colour) and by engaging in physical contact, implemented by the robot suggesting to do a fist-bump gesture, which could further increase the sense of belongingness [20].

The results provide much-needed evidence that it is possible to overcome the wearing off of motivation over time, which has been reported in various studies using social robots for learning tasks [7, 19, 22] due to the novelty effect [19]. And since, within an educational setting, long-term interactions with a social robot are essential, children must remain engaged over time. That said, children in this experiment only interacted with the robot once after a group introduction. Further research is needed to investigate the effects of our strategies in long-term settings.

The experimental design was such that all needs described in SDT were supported by multiple strategies (e.g. granting a choice, naming the child by name, do a fist-bump gesture). As a result, it is impossible to know exactly to what extent each strategy contributed to our findings. Future research should investigate which strategies of our intervention boosted engagement. In addition, our task design was such that the task children had to fulfil was a monotonous one (30 rounds of *I spy with my little eye*), which would likely reduce motivation in itself. While we expected that choosing this task could reveal more clearly the effects of our intervention, a real-world application should consider a less monotonous task. For example by adapting the task in varying the images. Furthermore, the spatial arrangement of the robot could have an impact on annotating engagement. In our design the robot was placed at a 45-degrees angle as a motivational intervention to satisfy the need for relatedness, however Johal et al. found a higher engagement in a face-to-face setting, which they compared to a side-by-side setting [13]. Future research is needed to isolate the effect of spatial arrangement on engagement.

Contrary to our expectations, our study did not find a significant difference in immediate nor delayed learning gain between the three conditions. Given the relatively large differences in learning gain between a robot that provided a feeling of autonomy, competence and relatedness compared to the two other conditions (see Figure 6), it was somewhat surprising to find that this difference was not significant. However, these findings are in line with various other studies investigating the effect of motivation on learning that also did not find a significant effect [16, 25, 31]. A possible reason for this could be the large variations observed in learning gains indicating large individual differences between children. Such individual differences could be attributed to children's different developmental backgrounds, cognitive capacities, attention span, or other factors. Together with a lack of statistical power, a significant effect is then hard to attain. A power analysis demonstrates that, to observe a significant difference ($\alpha = .05$) with a moderate effect ($f = 0.25$) and a power of 0.95, we would need to have a sample size of over 250 children. We expect that with more power, it would have been likely that the *ACR-condition* would have revealed a significant effect. We subscribe the need for a well-powered study to detect a true effect in future research. In addition, it appeared that many children already knew the English word ladybug from a television program starring a ladybug, and that most children knew the English word monkey from a famous indoor playground in the Netherlands. As a result, the potential learning gain was lower for these children, which may have affected the findings.

Of course, it could be possible that there are really no effects on learning gain. If this is the case, we should consider alternative explanations. One explanation might be the that the duration of the experiment was too short to obtain a difference in learning gain. Perhaps children who are motivated on all three needs could experience increased intrinsic motivation for a longer period of time, which could improve their learning performance. A further study with multiple sessions is therefore recommended.

One reason for hypothesizing that our strategies to support intrinsic motivation and thus promote learning gains were based on educational human-human studies that found a mediating effect of engagement on learning performance [5]. However, again contrary to our predictions, we did not find convincing evidence that task engagement is strongly correlated with learning gain. We only found a positive correlation between task engagement and learning gain in the *AC-condition*. Again, this is likely due to the lack of power in the current study.

So, supporting children's intrinsic motivation using strategies to fulfil the basic needs for autonomy, competence and relatedness seems to have a positive effect on task engagement, and –if well powered– might promote learning gain (though we have no evidence for this latter effect). However, it remains unclear what aspect of the strategies that we implemented (autonomy, competence, relatedness) contributes to this. Would relatedness be sufficient? And if so, which aspect of relatedness (naming the child, distance to the robot, personalised colouring of the tablet background, the fist bump) has the desired effect? More experimental research is needed to investigate the individual effects of satisfying the needs for autonomy, competence, and relatedness, and assess for instance self-efficacy as marker of satisfying the need for competence.

6 CONCLUSION

In this study, we aimed to investigate how a social robot supporting basic human needs to stimulate their intrinsic motivation based on the Self-Determination Theory from psychology [8, 21] could affect children's task engagement and learning gain in second language tutoring. We found support for H1b in that a robot fulfilling the need for autonomy, competence and relatedness increases children's task engagement compared to the two conditions that did not support all these needs. Additionally, we found support for H2a in that when, at least some, psychological needs were supported children maintain more engaged towards the task than without any of these needs. It thus seems that fulfilling the need for relatedness, in addition to autonomy and competence, gave a boost to children's task engagement. This is important, since it shows that fulfilling all these needs seems crucial in maintaining higher levels of engagement, which in turn could lead to higher learning gains.

Although a higher level of learning gain was observed, we did not find evidence to support hypotheses H3 and H4 that children would learn and memorize more words when two or all three needs were supported. Moreover, we did not find support for hypothesis H5 that the levels of engagement correlate with learning gain. While these hypotheses were not confirmed, we do believe that with a better powered study –possibly combined with a long-term educational setting – the implementation of a social robot that supports the needs for autonomy, competence and relatedness, as suggested by the SDT theory, could yield promising results.

To conclude, this study suggests that it is beneficial to implement motivational strategies in a robot tutor based on the Self-Determination Theory [8, 23], supporting the psychological needs for autonomy, competence, and relatedness with promising results on children's task engagement and keeping children engaged over time. In particular, the need for relatedness seems positively affected children's motivation towards an L2 word learning task.

ACKNOWLEDGMENTS

The authors would like to thank Mirjam de Haas for providing her scheme and training to code task engagement, all reviewers for their valuable feedback, and all participating children, their parents, and the schools for their contribution to this experiment. The first author was supported by a research trainee grant provided by Tilburg University.

REFERENCES

[1] Minoo Alemi, Ali Meghdari, and Maryam Ghazisaedy. 2014. Employing humanoid robots for teaching English language in Iranian junior high-schools. *International Journal of Humanoid Robotics* 11, 03 (2014), 1450022.

[2] Minoo Alemi, Ali Meghdari, and Maryam Ghazisaedy. 2015. The impact of social robotics on L2 learners' anxiety and attitude in English vocabulary acquisition. *International Journal of Social Robotics* 7, 4 (2015), 523–535.

[3] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. 2018. Social robots for education: A review. *Science Robotics* 3, 21 (2018), eaat5954.

[4] Tony Belpaeme, Paul Vogt, Rianne Van den Berghe, Kirsten Bergmann, Tilbe Göksun, Mirjam De Haas, Junko Kanero, James Kennedy, Aylin C Küntay, Ora Oudgenoeg-Paz, et al. 2018. Guidelines for designing social robots as second language tutors. *International Journal of Social Robotics* 10, 3 (2018), 325–341.

[5] Phyllis C Blumenfeld, Toni M Kempler, and Joseph S Krajcik. 2006. *Motivation and cognitive engagement in learning environments.* na.

[6] Ginevra Castellano, Ana Paiva, Arvid Kappas, Ruth Aylett, Helen Hastie, Wolmet Barendregt, Fernando Nabais, and Susan Bull. 2013. Towards empathic virtual and robotic tutors. In *International Conference on Artificial Intelligence in Education.* Springer, 733–736.

[7] Jan de Wit, Thorsten Schodde, Bram Willemsen, Kirsten Bergmann, Mirjam de Haas, Stefan Kopp, Emiel Krahmer, and Paul Vogt. 2018. The effect of a robot's gestures and adaptive tutoring on children's acquisition of second language vocabularies. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 50–58.

[8] Edward L Deci and Richard M Ryan. 2000. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry* 11, 4 (2000), 227–268.

[9] Marina Fridin. 2014. Kindergarten social assistive robot: First meeting and ethical issues. *Computers in Human Behavior* 30 (2014), 262–272.

[10] Marina Fridin. 2014. Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education. *Computers & Education* 70 (2014), 53–64.

[11] Jeong-Hye Han, Mi-Heon Jo, Vicki Jones, and Jun-H Jo. 2008. Comparative study on the educational use of home robots for children. *Journal of Information Processing Systems* 4, 4 (2008), 159–168.

[12] Andri Ioannou, Emily Andreou, and Maria Christofi. 2015. Pre-schoolers' interest and caring behaviour around a humanoid robot. *TechTrends* 59, 2 (2015), 23–26.

[13] Wafa Johal, Alexis Jacq, Ana Paiva, and Pierre Dillenbourg. 2016. Child-robot spatial arrangement in a learning by teaching activity. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).* IEEE, 533–538.

[14] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. 2004. Interactive robots as social partners and peer tutors for children: A field trial. *Human–Computer Interaction* 19, 1-2 (2004), 61–84.

[15] James Kennedy, Paul Baxter, and Tony Belpaeme. 2015. The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* IEEE, 67–74.

[16] James Kennedy, Paul Baxter, Emmanuel Senft, and Tony Belpaeme. 2016. Social robot tutoring for child second language learning. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction.* IEEE Press, 231–238.

[17] James Kennedy, Séverin Lemaignan, Caroline Montassier, Pauline Lavalade, Bahar Irfan, Fotios Papadopoulos, Emmanuel Senft, and Tony Belpaeme. 2017. Child speech recognition in human-robot interaction: evaluations and recommendations. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 82–90.

[18] Terry K Koo and Mae Y Li. 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine* 15, 2 (2016), 155–163.

[19] Iolanda Leite, Carlos Martinho, and Ana Paiva. 2013. Social robots for long-term interaction: a survey. *International Journal of Social Robotics* 5, 2 (2013), 291–308.

[20] Kayako Nakagawa, Masahiro Shiomi, Kazuhiko Shinozawa, Reo Matsumura, Hiroshi Ishiguro, and Norihiro Hagita. 2011. Effect of robot's active touch on people's motivation. In *Proceedings of the 6th International Conference on Human-Robot Interaction.* ACM, 465–472.

[21] Christopher P Niemiec and Richard M Ryan. 2009. Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education* 7, 2 (2009), 133–144.

[22] Emmy Rintjema, Rianne van den Berghe, Anne Kessels, Jan de Wit, and Paul Vogt. 2018. A robot teaching young children a second language: The effect of multiple interactions on engagement and performance. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 219–220.

[23] Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology* 25, 1 (2000), 54–67.

[24] Eline Sierens, Maarten Vansteenkiste, Luc Goossens, Bart Soenens, and Filip Dochy. 2009. The synergistic relationship of perceived autonomy support and structure in the prediction of self-regulated learning. *British Journal of Educational Psychology* 79, 1 (2009), 57–68.

[25] Rianne van den Berghe, Sanne van der Ven, Josje Verhagen, Ora Oudgenoeg-Paz, Fotios Papadopoulos, and Paul Leseman. 2018. Investigating the effects of a robot peer on L2 word learning. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 267–268.

[26] Rianne van den Berghe, Josje Verhagen, Ora Oudgenoeg-Paz, Sanne van der Ven, and Paul Leseman. 2019. Social robots for language learning: A review. *Review of Educational Research* 89, 2 (2019), 259–295.

[27] Kurt VanLehn. 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist* 46, 4 (2011), 197–221.

[28] Maarten Vansteenkiste, Eline Sierens, Bart Soenens, Koen Luyckx, and Willy Lens. 2009. Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology* 101, 3 (2009), 671.

[29] Paul Vogt, Mirjam De Haas, Chiara De Jong, Peta Baxter, and Emiel Krahmer. 2017. Child-robot interactions for second language tutoring to preschool children. *Frontiers in Human Neuroscience* 11 (2017), 73.

[30] LS Vygotsky and Michael Cole. 1978. *Mind in Society.* Harvard University Press.

[31] Jacqueline Kory Westlund, Leah Dickens, Sooyeon Jeong, Paul Harris, David DeSteno, and Cynthia Breazeal. 2015. A comparison of children learning new words from robots, tablets, & people. In *Proceedings of the 1st International Conference on Social Robots in Therapy and Education.*

[32] Jacqueline M Kory Westlund, Marayna Martinez, Maryam Archie, Madhurima Das, and Cynthia Breazeal. 2016. Effects of framing a robot as a social agent or as a machine on children's social behavior. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).* IEEE, 688–693.

[33] Janie H Wilson and Lawrence Locker Jr. 2008. Immediacy scale represents four factors: Nonverbal and verbal components predict student outcomes. *Journal of Classroom Interaction* 42, 2 (2008), 4–10.