# Gender Differences in Learning-by-Teaching a Social Robot: Insights from a Primary School Study.

TARAKLI, Imene, VINANZI, Samuele and DI NUOVO, Alessandro <http://orcid.org/0000-0003-2677-2650>

# Gender Differences in Learning-by-Teaching a Social Robot: Insights from a Primary School Study

Imene Tarakli[1], Samuele Vinanzi [1], Alessandro Di Nuovo [1]

***Abstract*— Social robots hold great promise for supporting children's learning, yet their effectiveness may depend on how well they align with individual learner characteristics. This study investigates the role of gender in shaping the outcomes of Learning-by-Teaching (LbT) with a social robot. In a primary school setting, 53 children (aged 8–9) participated in French language tasks under either a robot-assisted LbT condition or a self-practice condition. While no significant effects were found, girls consistently showed higher learning and retention gains when teaching the robot compared to practicing alone, an effect not observed in boys. Contrary to expectations, girls and boys spent similar time on task, used help equally, and reported comparable perceptions of the learning activity. Exploratory analyses revealed that girls who found the task more difficult learnt more, aligning with theories of desirable difficulty, and that higher learning gains were linked to lower perceived competence, suggesting possible signs of Imposter Syndrome. These findings highlight the complex interplay between cognitive and emotional factors in robot-assisted learning and emphasise the need for personalised educational technologies that adapt not only to performance but also to learner identity and psychological experience.**

## I. INTRODUCTION

Social robots are increasingly recognised as valuable tools in education, offering unique benefits such as enhanced student motivation, sustained engagement, and improved learning outcomes [1], [2]. Their interactive and dynamic nature fosters deeper knowledge acquisition and more positive attitudes toward learning. While social robots already provide meaningful educational support, personalising their interactions can further amplify their effectiveness. Research in psychology [3], human-computer interaction [4], and human-robot interaction [5], [6], [7] suggests that tailoring educational experiences to individual learner characteristics enhances engagement, retention, and overall success.

Learning is not uniform; it is shaped by individual differences such as personality and gender, which influence how students engage with and benefit from educational experiences. Personality traits strongly predict academic performance. Conscientiousness, characterised by diligence and organisation, consistently correlates with higher achievement [8]. Openness and Extraversion also contribute positively, while Neuroticism makes learning more challenging. Gender further modulates learning styles and engagement. Studies suggest that boys tend to favour hands-on, visual learning [9], [10], [11], while girls privilege verbal and collaborative tasks [12], [13]. Additionally, girls typically demonstrate greater school engagement, effort, and attentiveness, often translating into higher academic success [14].

[1]Department of Computing, Sheffield Hallam University, Sheffield, UK. i.tarakli@shu.ac.uk

These gendered learning patterns also extend to interactions with social robots. Research shows that girls respond more positively to robotic peers [15] and maintain higher levels of anthropomorphism over time, which correlates with stronger retention [16]. This raises the question: do these differences also shape Learning-by-Teaching (LbT) outcomes when children instruct social robots?

LbT has well-documented benefits, it encourages deeper cognitive processing, sustains attention, and strengthens metacognitive skills by requiring students to evaluate both their own and their tutee's knowledge [17]. Central to this process is the *Protégé Effect*, a psychological phenomenon in which teachers, here children, develop a sense of responsibility for their student, leading to greater commitment and learning [18].

This study builds on our previous work, where we implemented Interactive Reinforcement Learning as a cognitive model for a peer-like robot in an LbT scenario with 58 primary school children. Our findings demonstrated that teaching a social robot led to significantly higher engagement and retention compared to self-practice. Given established gender differences in communication, social interaction, and motivation, we now investigate whether these findings are influenced by gender. Specifically, we explore whether girls' heightened empathy and social skills make them more receptive to the Protégé Effect, ultimately enhancing their learning gains through LbT.

Specifically, our research questions are:

- Do girls experience greater learning gains from LbT compared to self-practice, relative to boys?
- Do girls perceive the learning activity more positively when teaching a robot?

By addressing these questions, this study aims to deepen our understanding of how gender differences impact social robot-assisted learning and to shed light on how these differences can inform the design of more effective, inclusive, and personalised educational technologies. Ultimately, this work contributes to a growing body of research advocating for adaptive, learner-centered HRI systems, where personalisation is driven by an understanding of individual differences; an important step toward maximising the educational potential of social robots.

Fig. 1: A child interacts with the robot in the Learning-by-Teaching condition, guiding the robot's learning by giving evaluative feedback on its responses via a tablet.

## II. BACKGROUND

### A. Social robots in education

Social robots have gained considerable attention in education due to their embodied presence and ability to engage children through rich social interactions [2]. Unlike screen-based or virtual agents, robots can leverage nonverbal cues such as gestures and gaze to support learning [19], [20], [21]. Their physical embodiment fosters stronger emotional and cognitive engagement [22], [23], [24], enhances memory retention [5], [25], and promotes social facilitation effects; where performance improves in the presence of a social agent [26], [27]. Beyond their presence, robots can take on a variety of pedagogical roles, including tutor, peer, or novice learner [1]. One particularly promising role is that of a novice peer, where the robot is framed as a learner in need of help. This setting enables the LbT paradigm, where children adopt the role of a tutor, guiding the robot through instruction and feedback. LbT has been shown to enhance children's learning outcomes by triggering metacognitive processes, sustained attention, and deeper cognitive engagement [28], [29]. Moreover, novice robots elicit caretaking behaviours [30], reduce performance anxiety [28], and create a safe, collaborative learning space that supports children's confidence and motivation [31].

### B. Personalisation in Robot-assisted learning

Personalisation has been consistently shown to enhance both cognitive and affective outcomes in child–robot educational interactions. Several authors found that tailoring robot behavior to the individual learner leads to higher engagement, deeper learning, and stronger social rapport. For instance, Leyzberg et al. demonstrated that personalising a robot tutor's feedback style to match a learner's preferences improved task performance and user satisfaction [32]. Similarly, Brown and Howard found that personalised verbal encouragement not only boosted learning outcomes but also enhanced the emotional perception of the robot [33]. In another study, Schodde et al. showed that adaptive feedback based on real-time performance led to higher vocabulary retention and more positive user experience [34]. Similarly, Lubold et al. showed that prosodic adaptation and turn-taking adjustments increased rapport and focus during storytelling tasks [35]. These findings motivate us to consider the importance of understanding individual learner differences, such as gender-related variations in engagement, empathy, and communication style, as an important step toward designing robots that can adapt their behaviours to the specific needs of each student, which could further enhances learning outcomes.

## III. METHODOLOGY

This study builds on previous work, where we investigated the impact of LbT with a social robot on children's language learning. In this paper, we extend that work by exploring how gender may influence learning outcomes and perceptions within the same paradigm. This analysis is based on the same dataset as our prior study, but focuses specifically on gender-related effects. Below, we provide an overview of the methodology. For full technical and procedural details, including the overall learning effects and experimental design, we refer the reader to the original publication [36].

### A. Participants

A total of 53 children (33 boys, 16 girls, and 4 who preferred not to disclose their gender), aged 8–9 years, from a UK primary school participated in the study. All participants were monolingual English speakers learning French as a second language. The study was conducted in a real classroom environment, with multiple child–robot interactions taking place simultaneously. Because the current analysis focuses specifically on exploring gender-related differences in learning outcomes and perceptions, only data from children who self-identified as either male or female were included. As a result, the final sample for analysis comprised 49 children. This decision was made solely for analytical consistency in addressing the research questions and does not reflect any exclusionary intent. Ethical approval was obtained from the Sheffield Hallam University Research Ethics Committee (Application ID: ER65839196, approved on July 17, 2024).

### B. Study Design

The experiment followed a between-subjects design applied separately to two distinct tasks:

- *Body Parts Task (memorisation-based):* Children matched English body part names to their French equivalents (e.g., "head" $\rightarrow$ tête).
- *Grammar Task (rule-inference-based):* Children categorised French nouns as masculine, feminine, or plural based on their determiners (un, une, les).

For each task, children were randomly assigned to one of two conditions: (1) *Learning-by-Teaching (LbT)*, where the child acted as a tutor, giving feedback on the robot's answers, and (2) *Self-Practice*, where the child performed the same activity alone on a tablet, without the robot.

Each child participated in both tasks, but experienced the robot condition in only one, to avoid cross-condition effects. Figure 1 shows the setting of the study for the LbT condition.

### C. Robots & Cognitive Model

In the LbT conditions, the robot was framed as a novice peer and operated using a teachable Interactive Reinforcement Learning (IRL) model. We used the JD robot from EZ-Robot, a small humanoid robot with expressive LED eyes, articulated limbs, and built-in speech synthesis. Its child-friendly appearance and compact size made it well-suited for primary school environments and for fostering peer-like interaction.

The robot began with no prior knowledge and made random guesses. It then improved over time based on the binary feedback (correct/wrong) provided by the child on the tablet after each response. A Q-learning algorithm was used to update the robot's action policy:

- State space: Representing word categories (Body Parts or grammatical category)
- Action space: Representing possible answers (e.g., masculine, feminine, plural).

This adaptive behaviour allowed the robot to gradually demonstrate learning progress, reinforcing its role as a novice peer in need of guidance and promoting the child's engagement through the Learning-by-Teaching paradigm. While the robot's behaviour was not dynamically adapted based on gender, this was an intentional design choice to serve as a control parameter, allowing us to investigate how a uniform robot behaviour is perceived and responded to by learners of different genders. In both tasks, the robot responded verbally to the child's teaching attempts, and its answer was displayed on the tablet screen. The child then provided binary evaluative feedback (correct/wrong) by selecting the appropriate option. After the feedback was submitted, the correct answer was revealed on the tablet, allowing the child to compare it with the robot's response. At any point during the task, children could press a help button on the tablet, which revealed the correct answer. This support tool was identical across both conditions and tasks.

### D. Procedure & Measures

A group familiarisation session was held one week before the study to introduce children to the JD robot. This brief session allowed participants to observe the robot's appearance, speech, and movements, helping to reduce novelty effects and ensure comfort with the setup.

Each child completed a pre-test, a learning phase, a post-test, and a retention test that occurred two weeks later. All learning activities took place on a tablet interface, consistent across conditions. The following data were collected:

- **Learning performance:** Based on test scores across the three phases.
- **Interaction metrics:** Time per question, hint use, and feedback behaviour.
- **Perception measures:** Task enjoyment, perceived competence, perceived engagement and task difficulty, via adapted child-friendly Likert scales.

## IV. RESULTS

### A. Learning Gain

We explored whether the LbT condition led to greater learning gains compared to Self-Practice, and whether this effect differed by gender. Learning gains were calculated as the difference between post-test and pre-test scores, while retention gains were based on the difference between retention and pre-test scores. Figure 2 illustrates the learning gains for both tasks.

In the body parts task, girls in the LbT condition showed higher learning gains ($M = 2.80$, $SD = 5.71$) than those in the Self-Practice condition ($M = -1.00$, $SD = 1.73$). Although this difference was not statistically significant, $U = 54.50$, $z = 1.90$, $p = .056$, the medium effect size ($r = .46$) suggests a meaningful trend. A similar pattern was found for retention gain, with girls retaining more vocabulary in the LbT condition ($M = 1.30$, $SD = 4.79$) compared to Self-Practice ($M = -3.57$, $SD = 4.31$), $U = 55.00$, $z = 1.95$, $p = .055$, $r = .47$.

In contrast, boys performed similarly across both conditions. Their average learning gains were comparable (LbT: $M = 2.76$, $SD = 5.39$; Self-Practice: $M = 1.43$, $SD = 3.91$), with no significant difference, $U = 150.50$, $z = 0.52$, $p = .61$, $r = .09$. Retention scores also remained stable (LbT: $M = 0.70$, $SD = 3.44$; Self-Practice: $M = -0.88$, $SD = 4.86$), $U = 168.50$, $z = -0.91$, $p = .36$, $r = -.15$.

In the grammar task, girls again showed higher learning gains in the LbT condition ($M = 5.33$, $SD = 4.46$) compared to Self-Practice ($M = 2.00$, $SD = 4.10$), with a medium effect size ($r = .39$), though this difference was not significant, $U = 44.50$, $z = 1.57$, $p = .13$. Retention followed a similar trend (LbT: $M = 4.67$, $SD = 4.46$; Self-Practice: $M = 0.60$, $SD = 4.74$), $U = 42.50$, $z = 1.36$, $p = .17$, $r = .34$.

Boys' performance in the grammar game did not differ meaningfully across conditions. Learning gains were nearly identical (LbT: $M = 2.40$, $SD = 4.88$; Self-Practice: $M = 1.59$, $SD = 5.19$), $U = 128.00$, $z = 0.02$, $p = 1.00$, $r = .003$, and retention gains were similarly consistent (LbT: $M = 1.00$, $SD = 3.05$; Self-Practice: $M = 0.94$, $SD = 5.20$), $U = 140.00$, $z = 0.47$, $p = .65$, $r = .08$.

We also conducted exploratory comparisons of interaction patterns across genders and tasks (see Table I). No significant differences were found in answer accuracy, time per iteration, or time spent on help. An exception was observed in the grammar game, where boys in the Self-Practice condition spent significantly more time per iteration ($Mdn = 3082.60$ ms) than girls ($Mdn = 1234.90$ ms), $U = 137.00$, $z = -2.41$, $p = .02$, $r = -.46$.

### B. Perception of the learning activity

The perception of the learning activity was assessed across task enjoyment, perceived competence, engagement, and

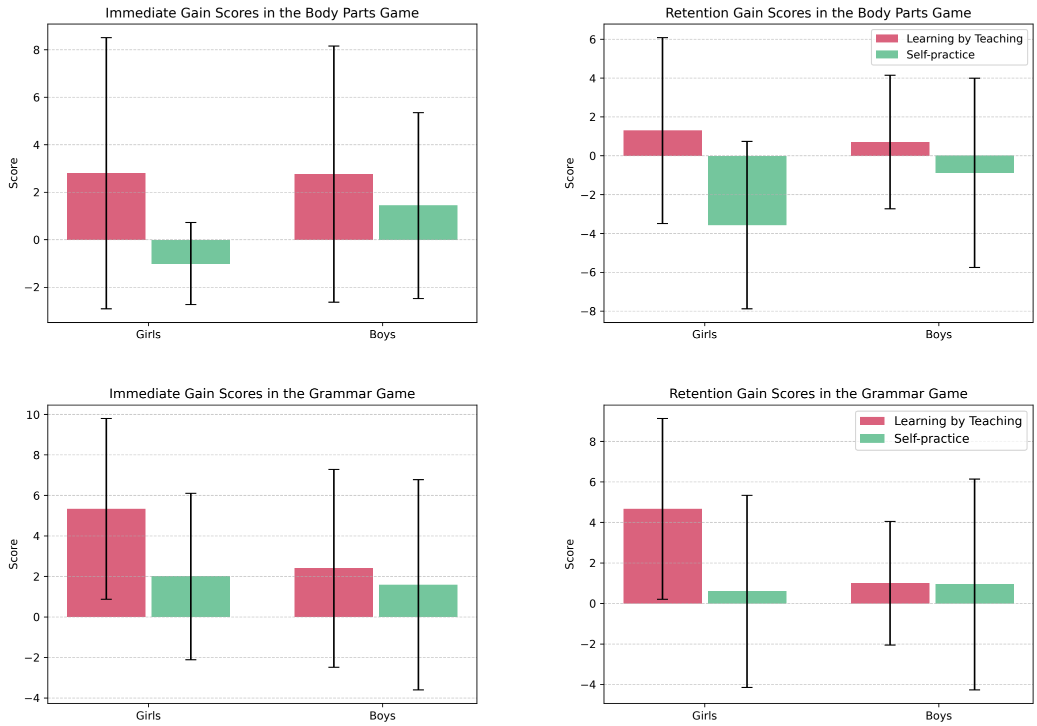


Fig. 2: Immediate and retention learning gains by gender and condition (LbT vs. Self-Practice) for both language tasks. Girls in the LbT condition consistently show higher learning and retention scores compared to other groups.

difficulty. Results are described in Table I. No significant differences were found between the LbT and Self-Practice conditions for both genders on any of these dimensions, in either the Body Parts or Grammar tasks.

For task enjoyment and perceived competence, ratings were generally consistent across conditions for both girls and boys, with no statistically significant differences. Engagement levels showed some variation, but again, no significant differences were observed. Boys tended to report higher engagement and perceived competence in the Grammar task, while girls generally rated the tasks similarly across conditions.

Difficulty ratings were comparable for both tasks, with no significant differences between conditions. While boys rated the tasks as somewhat easier, particularly in the Grammar game, the differences were not statistically significant.

### C. Exploratory analysis

Exploratory analyses revealed some notable correlations for girls in the Learning-by-Teaching (LbT) condition. A significant positive correlation was found between task difficulty and learning gain ($r = 0.553, p = 0.026$), indicating that as girls rated the task as more difficult, they showed higher learning gains. This correlation was not observed for girls in the Self-Practice condition or for boys in either condition.

Additionally, a strong negative correlation was found between perceived competence and learning gain for girls in the LbT condition ($r = -0.7, p = 0.002$), suggesting that as girls showed higher learning gains, they felt less competent. This pattern may reflect Imposter Syndrome, where individuals who perform well may feel undeserving of their success. This correlation was not observed for either boys or girls in the Self-Practice condition.

These findings suggest that for girls, the perceived difficulty of the task and their perceived competence may be important factors influencing learning outcomes in the LbT condition, while such correlations were not present for boys or in the Self-Practice condition.

## V. DISCUSSION & FUTURE WORKS

This analysis set out to explore the role of gender in shaping learning outcomes and perceptions in a Learning-by-Teaching (LbT) scenario with a social robot. Although our results did not reach statistical significance, we observed consistent trends suggesting that girls benefited more from the LbT condition compared to Self-Practice, while boys performed similarly across both conditions. These trends raise important questions about the mechanisms driving learning gains in socially embedded robot-assisted learning.

One might assume that higher learning gains would be explained by increased time spent on the learning activity, greater use of support tools (such as the help panel), or higher engagement with the task. However, our data showed no such differences between girls and boys. Both groups spent similar amounts of time, used help features at comparable rates, and showed similar levels of task enjoyment, engagement, and perceived competence. Interestingly, this lack of difference in observable engagement metrics challenges our initial expectation that girls' enhanced performance in the LbT condition would stem from greater behavioral involvement. Instead,

| Variable | Body Parts Game | | | | | | Grammar Game | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Learning by Teaching | | | Self-Practice | | | Learning by Teaching | | | Self-Practice | | |
| | Girls | Boys | p-value | Girls | Boys | p-value | Girls | Boys | p-value | Girls | Boys | p-value |
| Task Enjoyment | 4.60 | 4.60 | 1.0 | 4.60 | 4.50 | 0.65 | 4.60 | 4.80 | 0.87 | 4.60 | 4.20 | 0.42 |
| Perceived Competence | 4.20 | 4.40 | 0.88 | 4.40 | 4.50 | 0.71 | 3.70 | 5.00 | 0.30 | 4.60 | 4.60 | 0.84 |
| Engagement | 4.17 | 4.00 | 0.88 | 4.33 | 5.00 | 0.11 | 4.33 | 5.00 | 0.24 | 4.16 | 3.66 | 0.32 |
| Difficulty | 2.00 | 2.00 | 0.28 | 2.00 | 2.50 | 0.78 | 3.00 | 2.00 | 0.27 | 3.50 | 2.00 | 0.15 |

TABLE I: Comparison of LbT and Self-Practice conditions across engagement, competence, enjoyment, and difficulty for both learning tasks.

the findings suggest that the learning gains may have been driven by deeper cognitive or motivational mechanisms such as heightened metacognitive processing, a stronger sense of responsibility when teaching, or internalised motivation triggered by the social dynamic of the LbT condition.

If the effort and engagement levels were similar across genders, how can we explain the improved learning outcomes in girls? One possible explanation lies in individual learning preferences. Prior research has shown that girls tend to prefer verbal and socially collaborative learning contexts, which the LbT paradigm naturally supports. The act of teaching a robot, giving feedback, and reflecting on the robot's responses likely aligns more closely with girls' learning styles, leading to deeper cognitive processing. However, our exploratory analyses offer a more nuanced view. Girls in the LbT condition showed greater learning gains when they perceived the task as more difficult. This finding aligns with the concept of desirable difficulties , which posits that tasks that require greater mental effort can lead to better long-term retention and transfer of knowledge [37]. It is possible that the social responsibility of teaching the robot—known as the Protégé Effect—was stronger in girls, prompting them to invest more cognitive effort when they felt challenged. In turn, this deeper engagement with the material could have resulted in enhanced learning.

We found a strong negative correlation between post-test scores and perceived competence among girls, suggesting that those who performed better felt less confident. To rule out the possibility that this effect was driven by pre-existing differences in ability, we examined the correlation between pre-test scores and perceived competence and found no significant relationship. This might mirror patterns associated with Imposter Syndrome, a phenomenon where individuals, despite being successful, internalise self-doubt and feel undeserving of their achievements [38]. In this context, girls may have experienced the pressure of acting as the robot's "teacher," resulting in increased cognitive effort to meet that responsibility, and heightened interpersonal awareness [39].

While medium-to-large effect sizes were observed, our sample was underpowered for detecting interaction effects. These findings should be interpreted as preliminary and warrant replication with a larger and gender-balanced sample. Longitudinal studies are also needed to assess whether these gender-based interaction patterns persist over time and influence long-term learning outcomes. In addition, future research should incorporate qualitative data, such as interviews or interaction logs, to uncover the mechanisms underlying these gendered differences in response to Learning-by-Teaching.

## VI. CONCLUSIONS

This study highlights how gender may influence the effectiveness of LbT with social robots. While no significant differences were found, trends indicated that girls benefited more from LbT than from self-practice, whereas boys performed similarly across conditions. These patterns suggest that the social and verbal nature of LbT may align more closely with girls' learning preferences. Exploratory findings also revealed deeper psychological dynamics. Girls who perceived the task as more difficult achieved greater learning gains, consistent with the theory of desirable difficulties. Additionally, those who learned more reported lower perceived competence, suggesting signs of Imposter Syndrome, where success is accompanied by self-doubt. Together, these findings point to the importance of considering individual learner differences, not only in behavior but also in emotional and cognitive responses, when designing robot-assisted learning experiences. Future work should validate these trends in larger and more diverse samples and explore how robots might adapt to support both the cognitive and emotional needs of different learners.


## ACKNOWLEDGMENT

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955778 and was funded by the Innovate UK (grant number 10089807 for the Horizon Europe project PRIMI Grant agreement n. 101120727).For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.


## REFERENCES


[1] W. Johal, "Research trends in social robots for learning," *Current Robotics Reports*, vol. 1, no. 3, pp. 75–83, 2020.

[2] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, "Social robots for education: A review," *Science robotics*, vol. 3, no. 21, p. eaat5954, 2018.

[3] B. S. Bloom, "The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring," *Educational researcher*, vol. 13, no. 6, pp. 4–16, 1984.

[4] J. R. Anderson, C. F. Boyle, A. T. Corbett, and M. W. Lewis, "Cognitive modeling and intelligent tutoring," *Artificial intelligence*, vol. 42, no. 1, pp. 7–49, 1990.

[5] J. De Wit, T. Schodde, B. Willemsen, K. Bergmann, M. De Haas, S. Kopp, E. Krahmer, and P. Vogt, "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*, 2018, pp. 50–58.

[6] A. Ramachandran, S. S. Sebo, and B. Scassellati, "Personalized robot tutoring using the assistive tutor pomdp (at-pomdp)," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 8050–8057.

[7] X. Zhang, C. Breazeal, and H. W. Park, "A social robot reading partner for explorative guidance," in *proceedings of the 2023 ACM/IEEE international conference on human-robot interaction*, 2023, pp. 341–349.

[8] S. Mammadov, "Big five personality traits and academic performance: A meta-analysis," *Journal of personality*, vol. 90, no. 2, pp. 222–255, 2022.

[9] E. E. Maccoby and C. N. Jacklin, *The Psychology of Sex Differences:—Vol. II: Annotated Bibliography*. Stanford university press, 1978, vol. 2.

[10] M. C. Linn and A. C. Petersen, "Emergence and characterization of sex differences in spatial ability: A meta-analysis," *Child development*, pp. 1479–1498, 1985.

[11] G. Iaria, M. Petrides, A. Dagher, B. Pike, and V. D. Bohbot, "Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: variability and change with practice," *Journal of Neuroscience*, vol. 23, no. 13, pp. 5945–5952, 2003.

[12] K. Bolla-Wilson and M. L. Bleecker, "Influence of verbal intelligence, sex, age, and education on the rey auditory verbal learning test," *Developmental Neuropsychology*, vol. 2, no. 3, pp. 203–211, 1986.

[13] E. Capitani, M. Laiacona, and A. Basso, "Phonetically cued word-fluency, gender differences and aging: A reappraisal," *Cortex*, vol. 34, no. 5, pp. 779–783, 1998.

[14] S.-f. Lam, S. Jimerson, E. Kikas, C. Cefai, F. H. Veiga, B. Nelson, C. Hatzichristou, F. Polychroni, J. Basnett, R. Duck *et al.*, "Do girls and boys perceive themselves as equally engaged in school? the results of an international study from 12 countries," *Journal of school psychology*, vol. 50, no. 1, pp. 77–94, 2012.

[15] K. Balkibekov, S. Meiirbekov, N. Tazhigaliyeva, and A. Sandygulova, "Should robots win or lose? robot's losing playing strategy positively affects child learning," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2016, pp. 706–711.

[16] R. Van den Berghe, M. de Haas, O. Oudgenoeg-Paz, E. Krahmer, J. Verhagen, P. Vogt, B. Willemsen, J. de Wit, and P. Leseman, "A toy or a friend? children's anthropomorphic beliefs about robots and how these relate to second-language word learning," *Journal of Computer Assisted Learning*, vol. 37, no. 2, pp. 396–410, 2021.

[17] D. Duran, "Learning-by-teaching. evidence and implications as a pedagogical mechanism," *Innovations in education and teaching international*, vol. 54, no. 5, pp. 476–484, 2017.

[18] C. C. Chase, D. B. Chin, M. A. Oppezzo, and D. L. Schwartz, "Teachable agents and the protégé effect: Increasing the effort towards learning," *Journal of science education and technology*, vol. 18, pp. 334–352, 2009.

[19] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, "The physical presence of a robot tutor increases cognitive learning gains," in *Proceedings of the annual meeting of the cognitive science society*, vol. 34, no. 34, 2012.

[20] J. Li, "The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents," *International Journal of Human-Computer Studies*, vol. 77, pp. 23–37, 2015.

[21] J. Kennedy, P. Baxter, and T. Belpaeme, "The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning," in *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, 2015, pp. 67–74.

[22] A. Vrins, E. Pruss, J. Prinsen, C. Ceccato, and M. Alimardani, "Are you paying attention? the effect of embodied interaction with an adaptive robot tutor on user engagement and learning performance," in *International conference on social robotics*. Springer, 2022, pp. 135–145.

[23] M. de Haas, P. Vogt, R. van den Berghe, P. Leseman, O. Oudgenoeg-Paz, B. Willemsen, J. de Wit, and E. Krahmer, "Engagement in longitudinal child-robot language learning interactions: disentangling robot and task engagement," *International Journal of Child-Computer Interaction*, vol. 33, p. 100501, 2022.

[24] P. Van Minkelen, C. Gruson, P. Van Hees, M. Willems, J. De Wit, R. Aarts, J. Denissen, and P. Vogt, "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*, 2020, pp. 369–377.

[25] H. Chen, H. W. Park, and C. Breazeal, "Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement," *Computers & Education*, vol. 150, p. 103836, 2020.

[26] N. Riether, F. Hegel, B. Wrede, and G. Horstmann, "Social facilitation with social robots?" in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 2012, pp. 41–48.

[27] A. Cruz-Maya, F. Ferland, and A. Tapus, "Social facilitation in a game-like human-robot interaction using synthesized emotions and episodic memory," in *Social Robotics: 7th International Conference, ICSR 2015, Paris, France, October 26-30, 2015, Proceedings 7*. Springer, 2015, pp. 164–173.

[28] S. Lemaignan, A. Jacq, D. Hood, F. Garcia, A. Paiva, and P. Dillenbourg, "Learning by teaching a robot: The case of handwriting," *IEEE Robotics & Automation Magazine*, vol. 23, no. 2, pp. 56–66, 2016.

[29] E. Yadollahi, W. Johal, A. Paiva, and P. Dillenbourg, "When deictic gestures in a robot can harm child-robot collaboration," in *Proceedings of the 17th ACM conference on interaction design and children*, 2018, pp. 195–206.

[30] F. Tanaka and S. Matsuzoe, "Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning," *Journal of Human-Robot Interaction*, vol. 1, no. 1, pp. 78–95, 2012.

[31] A. Jacq, S. Lemaignan, F. Garcia, P. Dillenbourg, and A. Paiva, "Building successful long child-robot interactions in a learning context," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 239–246.

[32] D. Leyzberg, S. Spaulding, and B. Scassellati, "Personalizing robot tutors to individuals' learning differences," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 423–430.

[33] L. N. Brown and A. M. Howard, "The positive effects of verbal encouragement in mathematics education using a social robot," in *2014 IEEE integrated STEM education conference*. IEEE, 2014, pp. 1–5.

[34] T. Schodde, K. Bergmann, and S. Kopp, "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*, 2017, pp. 128–136.

[35] N. Lubold, E. Walker, H. Pon-Barry, and A. Ogan, "Comfort with robots influences rapport with a social, entraining teachable robot," in *Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part I 20*. Springer, 2019, pp. 231–243.

[36] I. Tarakli, Vinanzi, Samuele, Moore, Richard, and A. Di Nuovo, "Robots and children that learn together : Improving knowledge retention by teaching peer-like interactive robots," 2023.

[37] E. L. Bjork, R. A. Bjork *et al.*, "Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning," *Psychology and the real world: Essays illustrating fundamental contributions to society*, vol. 2, no. 59-68, 2011.

[38] P. R. Clance and S. A. Imes, "The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention." *Psychotherapy: Theory, research & practice*, vol. 15, no. 3, p. 241, 1978.

[39] B. A. Tewfik, "The impostor phenomenon revisited: Examining the relationship between workplace impostor thoughts and interpersonal effectiveness at work," *Academy of Management Journal*, vol. 65, no. 3, pp. 988–1018, 2022.