QWriter: A reinforcement learning-based robot for early literacy acquisition

Robot-assisted language learning produces comparable results to human tutors in a long-term study with elementary school children.

By Anara Sandygulova, Aida Amir, Nurziya Oralbayeva, Zhansaule Telisheva, Aida Zhanatkyzy, Aidar Shakerimov, Shamil Sarmonov, and Arna Aimysheva

DOI: 10.1145/3611679

Central Asian Kazakhstan has set out to return to the Kazakh Latin alphabet from the current Kazakh Cyrillic by 2031 [1]. The Cyrillic-to-Latin switchover may pose critical challenges to early literacy acquisition and development in children who need to learn to write in the two alphabets, Cyrillic and Latin, at the same time. Since most children's books and learning content remain predominantly available in the Kazakh Cyrillic script, learning the new alphabet poses a risk due to the inadequate educational and entertaining content that will be available in the new script. This may result in unwanted confusion and lead to decreased interest in using the Latin-based Kazakh alphabet. Our work aims to address these literacy-based challenges and proposes engaging learning applications for the acquisition of the Kazakh Latin alphabet and its handwriting for children in the primary grades.

A wealth of research demonstrates the advantages of learning a language for personal, social, and professional aspirations. With the rise of artificial tutoring systems, language learning has become affordable and multi-modal, adopting educational approaches while supporting technology-enhanced learning (e.g., learning by teaching in CoWriter). This motivation has driven the field of human-robot interaction (HRI) to shape robot-assisted language learning (RALL) as an evolving area of social learning. Social robots are commonly used to support language learning and literacy development. Compared to other technologies, robots can play a unique social role, for example as a peer, and have a physical embodiment [2]. They can help children practice language skills such as vocabulary knowledge, speaking, and reading comprehension

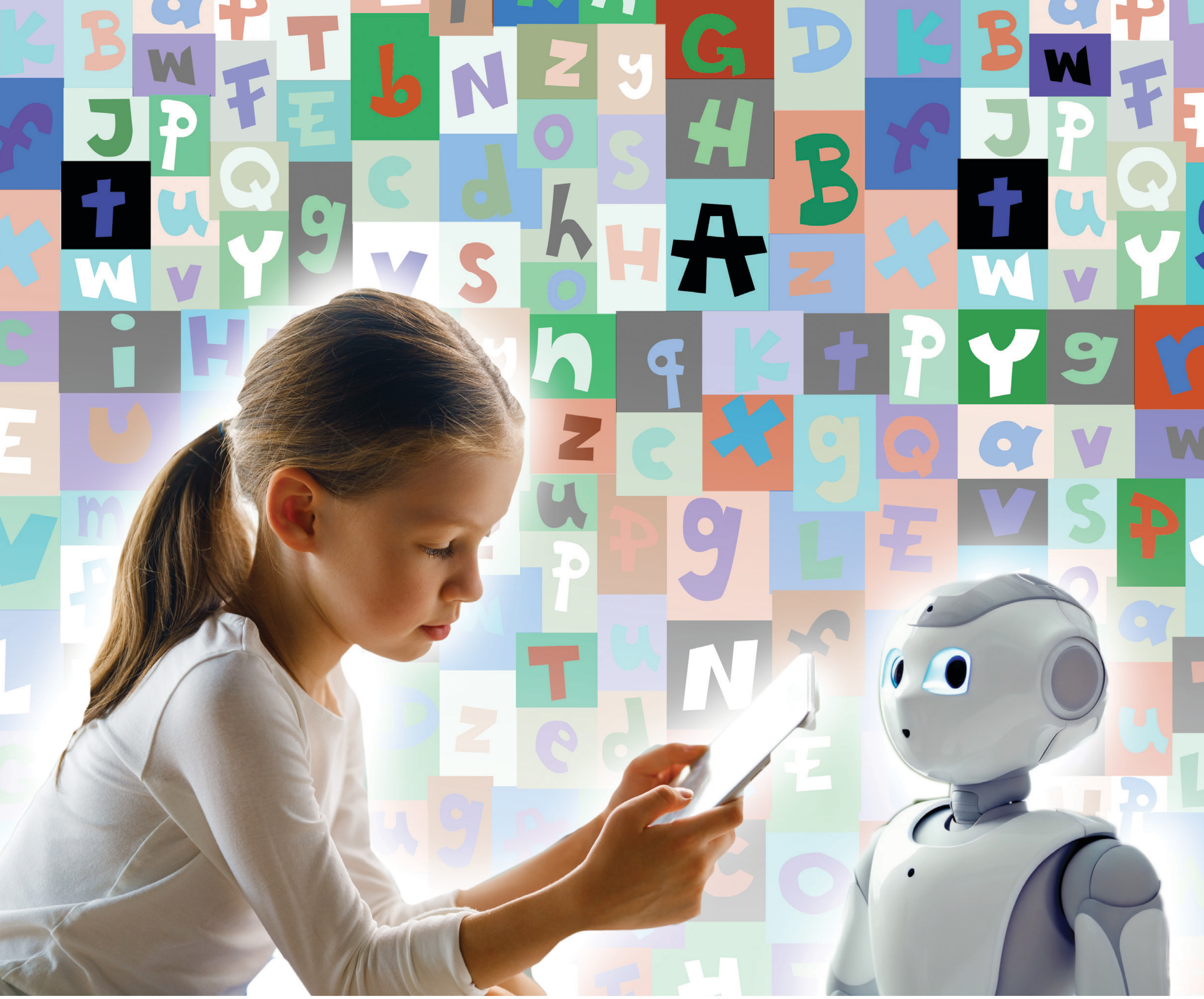


Photo by YuganovKonstantin / Shutterstock.com

[3]. This article contributes to the current body of knowledge about robots for language learning and proposes the RAAL environment to maintain learner engagement with an adaptive robot system.

Central to our work is reinforcement learning (RL), which supports adaptive personalization of HRI. It is a framework for decision-making algorithms in which the agent chooses actions when interacting with its environment and discovering optimal behavior [4]. RL differs from other machine learning algorithms, such as supervised and unsupervised learning, by its goal-directed learning from the environment that allows for a robust and adaptive interaction customized to a user's needs [5]. Currently, RL is gaining increasing scholarly attention as a way of personalization of HRI [4]. RL has gained recognition not only in the area of machine learning but also beyond, expanding its scope to language learning and other educational tasks and through long-term affective personalization [5].

This article investigates what happens when an RL-based robot system is implemented to maximize children's knowledge of the Kazakh Latin alphabet and its handwriting over repeated interactions, while adapting to each child's individual needs. We hypothesize that children can achieve similar rates of learning gains with the QWriter robot as with a human tutor. In the current study, the human tutor condition offers a baseline for what we consider a technology-enhanced way of learning in the modern world, such as the use of a tablet by a teacher. This understanding of social learning aspects of HRI, supported by RL algorithms, would contribute to the application of adaptive learning systems, addressing real-world challenges such as early literacy development amidst the alphabet switchover. To the best of our knowledge, no previous studies have investigated the RL-based robot system along with the human baseline; both applied in a public school.

THE QWRITER SYSTEM

In this work, we employ social robot NAO developed by SoftBank Robotics. It is one of the most popular commercially available social robots with an autonomous and programmable architecture. It has been widely used in HRI for robot-assisted applications in

Figure 1. A child misspells the letter in a word suggested by the QWriter system (left). The QWriter robot gives feedback by showing the correct spelling (right).

education and healthcare. The robot is 58 cm tall and portable. It is non-threatening and its human-like appearance is of great appeal to young users. It is equipped with seven touch sensors and has 25 degrees of freedom across various parts of its body. The robot's voice is a custom-made Kazakh speech synthesis. The robot is programmed to produce verbal and non-verbal social cues such as hand gestures, body behaviors, and speech utterances for a 30-minute activity. Another hardware component of the system includes the Wacom Cintiq Pro tablet, which has a stylus with 8.192 levels of pressure sensitivity and tilting recognition.

Learning Tasks. The alphabet learning activity includes 12 Kazakh words in each session, selected out of 60 in the overall word list presenting all the letters of the Kazakh Latin alphabet at least once. The robot "writes" on a Wacom Cintiq Pro tablet one letter at a time. A child writes the letters by copying the robot's writing in the first practice. Overall, there were five learning tasks across five sessions. We developed multi-stage learning tasks to develop basic literacy skills in the Kazakh Latin alphabet.

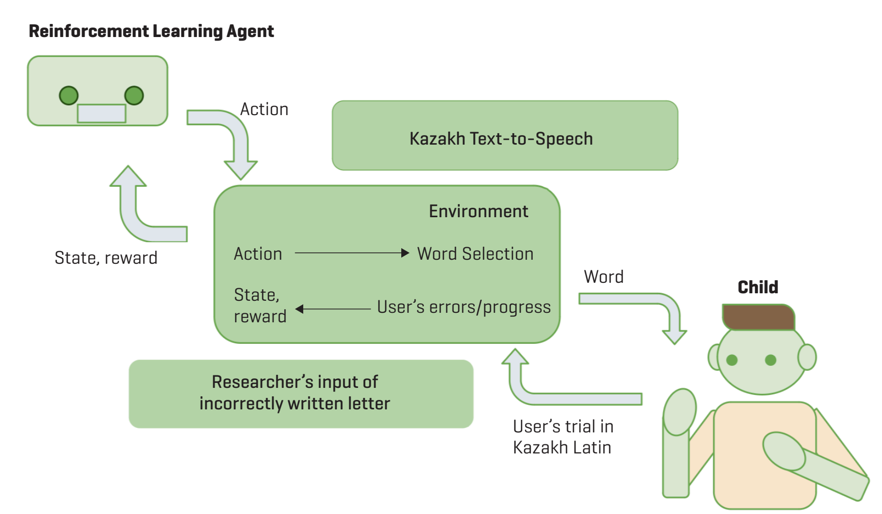
1. **Spelling bee.** A spelling bee is a spelling task that familiarizes children with the Kazakh Latin letters in a simple way. A child is asked to write a word in Kazakh Latin by copying what a robot imitates to write on the tablet (see Figure 1). The robot also pronounces the word and asks the child to repeat it. If the child misspells the letters, the next practice word is selected to reflect the child's mistake. For example, if the word "*salem*" is misspelled, the robot offers corrective feedback and writes the letter correctly. The next word is then selected in relation to this mistake so that the child has an opportunity to practice this letter once again. The system does not repeat words, choosing a new set of words for each session.

2. **Guess what?** This activity supports a letter practice that focuses on cognitive skills and visual perceptions. First, the robot writes a word with one letter missing. The robot explains the child has to guess a missing letter and spell the whole word. If the word is misspelled, the robot offers corrective feedback and writes the word again. When a child writes the word correctly, the robot expresses verbal encouragement and asks to continue.

3. **Scramble me.** This is a word-building game in which a child builds a new word using letters shown on the screen in random order. This activity enhances the attention span necessary for effective language learning. The robot writes a set of random letters on the screen and asks the child to guess and write the word. The children have to guess the right order of letters and write the word on the tablet. In case of wrong spelling, the child receives corrective feedback.

4. **Pic me.** This activity enables children to spell words without any written hints, however visual hints such as a picture describing a specific word can help support a child's engagement in the task. The child

Figure 2. The QWriter system environment.



sees pictures one by one. The robot only guides the child during the first try and then simply observes them by asking “What do you see in the next picture?” after they finish spelling words one at a time. Children have to read the words as they practice.

5. **Converter.** This activity results in the final stage in alphabet practice in which a child converts a word from Cyrillic to Latin in the Kazakh language. The robot displays a word written in Kazakh Cyrillic and asks the child to write it in Kazakh Latin. The robot explains the child has to convert and spell it in Kazakh Latin. The robot reminds the child to take time and retry.

Reinforcement Learning. Our RL-based QWriter system consists of an agent, environment, and user interface (see Figure 2). The agent is an RL algorithm that learns action selection from interaction with the environment. The environment represents an intermediate object that connects the agent with a user by recording information about the user’s mistakes and progress before transmitting it to the agent, as well as receiving an action from the agent and projecting it on interaction with the user. As a platform for user interaction, the humanoid robot NAO is used.

Environment. The environment works with the writing of words. Inside the environment is a set of variables that we considered to be useful in the learning process. They include lists of common Latin letters and specific Kazakh letters that a child has to learn, as well as lists of common Latin letters and specific Kazakh letters that are problematic for

A wealth of research demonstrates the advantages of learning a language for personal, social, and professional aspirations.

the child (i.e., the child writes it with error), a set of available words, and the currently assigned word.

The state of the agent is formulated from observations of the number of letters in the lists of unexplored and problematic letters. However, in order to keep the state-space small enough for faster exploration, we quantize the number of items in the lists in the following ways:

▸ Number of unexplored common Latin letters are divided into three groups: a) zero letters, b) more than zero but less than six, and c) more than six letters in the list.

▸ Number of unexplored specific Kazakh letters are divided into three groups: a) zero letters, b) more than zero but less than six, and c) more than six letters in the list.

▸ Number of problematic common Latin letters are divided into three groups: a) zero letters, b) more than zero but less than six, and c) more than six letters in the list.

▸ Number of problematic specific Kazakh letters are divided into three groups: a) zero letters, b) more than zero but less than six, and c) more than six letters in the list.

The state then is a tuple of four elements including the group index of number of unexplored common Latin letters (L), the group index of number of unexplored specific Kazakh letters (K), the group index of number of problematic common Latin letters (PL), and the group index of number of problematic specific Kazakh letters (PK).

Interaction with the user. The agent selects a word from the set of words and traces the letters’ writing on a tablet letter by letter. The robot then asks the child to copy the writing of the word with the help of the stylus. The letters contained in the word are removed from the list of unexplored letters. After recognizing the user’s spelling, the environment records errors. The resultant state is returned to the agent.

The agent selects an action or takes a random action in accordance with Epsilon-Greedy policy (i.e., with a probability ϵ it takes a random action instead of the best learned one). The options of action are:

▸ Select a word that contains unexplored common letters (if the corresponding list is not empty).

▸ Select a word that contains unexplored specific Kazakh letters (if the corresponding list is not empty).

▸ Select a word that contains a problematic common letter (if the corresponding list is not empty).

▸ Select a word that contains a problematic specific Kazakh letter (if

Figure 3. Experimental condition setups: human tutor (left) and QWriter robot (right).

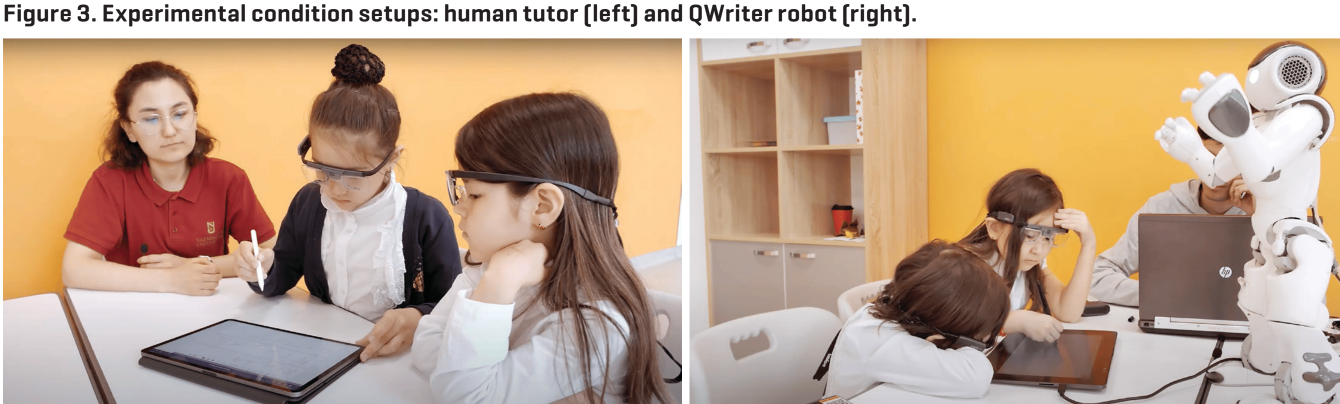
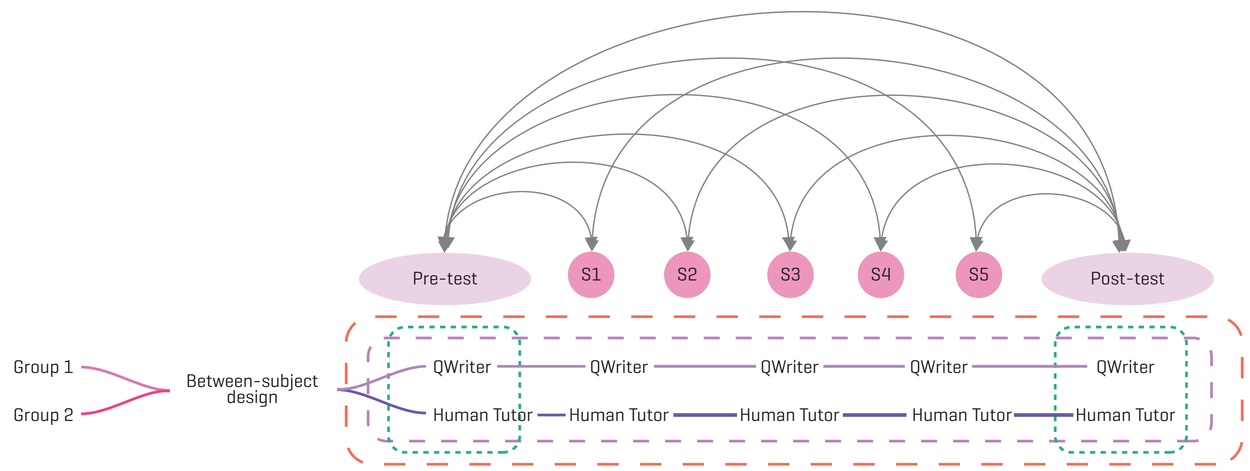


Figure 4. The experiment procedures.



the corresponding list is not empty).

The environment presents a new word to the user in accordance with the given action selection and records the child's errors. The obtained data from user interaction is returned as a new state to the agent.

At the end of 12 iterations, we provide a reward by calculating the number of encountered challenges (both solved and unsolved errors), the number of unsolved errors (if a child kept writing a certain letter with an error), and the number of unexplored letters.

OUR METHODS

This work was approved by the Institutional Research Ethics Committee at Nazarbayev University.

Recruitment. During recruitment in the classroom, children received assent and informed consent forms in the presence of their teachers. The forms included brief information including the aim of the research and its procedures. Once children returned the forms approved and signed by their parents, teachers collected them for us.

Participants. A total of 69 children aged 7–10 years old took part in the experiment. Due to the missing data of some children, the analysis for each session included a different number of children. The participants interacted with either a QWriter robot or a human tutor for five sessions on separate days, each lasting 30–40 minutes. Each interaction occurred in pairs (see Figure 3). We carefully counterbalanced children by their age and gender to avoid sampling bias.

Conditions. In this research, children learned in both conditions:

▸ QWriter robot (N=23); 12 words constitute the overall pool of 60 words, the order of which is chosen based on each child's mistakes. The letter mistakes are manually entered by the researcher using the keyboard on a laptop connected to the Wacom tablet. Along with the corrective feedback, the robot also provided auditory feedback to emphasize whether the child spelled the words correctly or not (e.g., "You got it right," "Well done," "You wrote letter X incorrectly," "Try your best next time.").

▸ Human Tutor (N=19); all 60 words are the same as in the QWriter condition, with the only difference being that the order of 12 words is fixed in each session. The teacher also provided feedback and corrected mistakes based on each child's individual mistakes (e.g., "Let me show you how to write this letter." "Now please repeat the letter by looking at my example."). The human tutor also used Kazakh for communication.

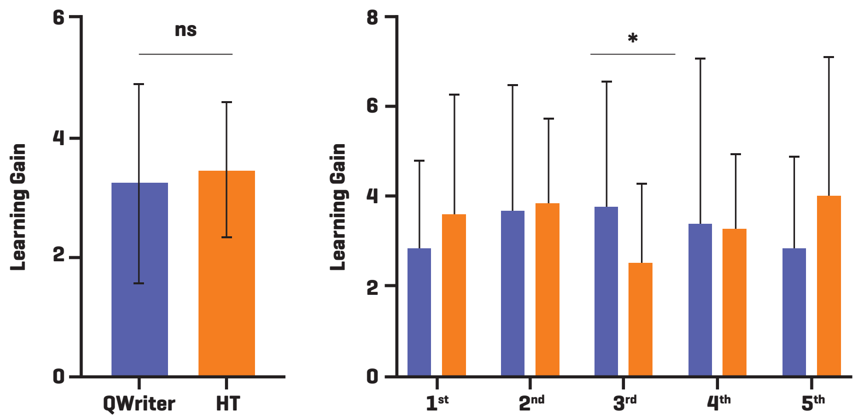
Procedure. The experiment consisted of a pre-test, a learning task, and a post-test in each session. Prior to the first session, a researcher conducted a survey asking demographic questions such as age, linguistic background, and academic performance. Children were randomly assigned to take a seat at the table with the robot or with the human tutor while counterbalancing the order of the conditions to avoid ordering effects. Before each meeting, children were asked to complete a pre-test to assess their knowledge of the Kazakh Latin alphabet. Each child wrote the test on a separate table. After each meeting, children completed a post-test to track session-specific learning gains. More details can be found in Figure 4.

Measures. Learning gain is the number of learned letters counted as a difference from a pre-test to a post-test that children did before and after each session with either the robot or the human. The test had two columns: one with Kazakh Cyrillic letters written in advance, and the other was blank to be filled in with Kazakh Latin letters by a child.

RESULTS

Our findings suggest overall learning gains between the two conditions were neither significantly different nor significantly equivalent as depicted in Figure 5 on the left. When analyzing individual sessions, we found a significant difference in learning gains in the third session between the conditions as depicted in Figure 5 on the right. The learning gains of the children in QWriter condition was 3.76 ± 2.79 letters on average, which is significantly higher than those in the human tutor condition (2.5 ± 1.78 letters). Other session-based learning

Figure 5. Average learning gains in both conditions (left), average learning gains by sessions (right).



gains did not yield statistically significant results.

DISCUSSION

The QWriter system can be as helpful as the human teacher in alphabet teaching. The difference in learning gains between the two conditions was neither significant nor equivalent, which suggests the QWriter provides more or less effective support as a human tutor. We accept our hypothesis. When analyzing sessional learning gains, particularly by S3, the children learned more letters with the QWriter robot than those children with the human tutor. The remaining sessions brought more or less similar learning outcomes in both conditions. The implemented system could offer some significant advantages to children. Its RL-specific trial-error-reward approach allowed the robot to discriminate between correct or wrong letters with corresponding feedback that allowed children to learn from their mistakes, which maximized learning gains. It is worth noting the human tutor did not adopt a traditional teaching role but rather maintained a learner-centered approach supported by a tablet with a stylus (an uncommon classroom tool). This fact is especially important as children in our context are usually taught in conventional ways—high-level teacher authority and low-tech classrooms. This result highlights the added value of using social robots in the classroom where human teachers focus on diversifying educational content and instruction.

CONCLUSION AND FUTURE WORKS

This study evaluated the QWriter system integrating an RL algorithm, a social robot, and a tablet with a stylus, which supported children in the acquisition of the Kazakh Latin alphabet and its handwriting. We tested the system in a primary school where 69 children aged 7–10 interacted with the QWriter robot that attempts to maximize learning gains while learning about the children's learning state by identifying misspelled letters and reintroducing them in the following practice words. The study shows the alphabet learning experiences of children were not significantly different between a human tutor and the robot tutor and the QWriter robot could be applicable in such a teaching scenario. The main contribution of the paper is an adaptive robot system using RL to learn an optimal policy for alphabet practice in child-robot interaction. The second contribution is the provision of empirical insights and the real-world application of the RL-based agent. These results provide an intermediary step for the HRI community to build more powerful educational systems.

ACKNOWLEDGMENT

This work was supported by the Nazarbayev University Collaborative Research Program. Grant numbers: 091019CRP2107 and OPCRP2022002.

References

[1] Fazylzhanova, A. Problems and prospects of reforming the Kazakh language on the Basis of the new Latin alphabet. Socio-political Aspects of the Realization of Kazakh Alphabet Switch to Latin. (2017), 191; http://iph.kz/doc/ru/1252.pdf

[2] van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., van der Ven, S. H. G., and Leseman, P. P. M. Social robots for language learning: A Review. *Review of Educational Research* 89, 2 (2018), 25–295; https://doi.org/10.3102/0034654318821286

[3] Chen, H., Park, H.W., and Breazeal, C. Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. *Computers & Education* 150, (2020); https://doi.org/10.1016/j.compedu.2020.103836

[4] Akalin, N. and Loutfi, A. Reinforcement learning approaches in social robotics. *Sensors* 21, 4 (2021); https://doi.org/10.3390/s21041292

[5] Shawky, D. and Badawi, A. Towards a personalized learning experience using reinforcement learning. In Hassanien, A. (eds.) *Machine Learning Paradigms: Theory and Application.* Studies in Computational Intelligence, vol 801. Springer, Cham. https://doi.org/10.1007/978-3-030-02357-7_8

Biographies

Anara Sandygulova is an associate professor in the Robotics and Mechatronics Department at Nazarbayev University, Astana, Kazakhstan. With a Ph.D. in computer science and a passion for leveraging technology, Sandygulova 's work focuses on creating adaptive systems integrating humanoid robots and machine learning to enhance educational experiences for young learners.

Aida Amir holds an M.A. in multilingual education from Nazarbayev University's Graduate School of Education. Her current research interests include bi/multilingual education, educational research, and language learning and autism in human-robot interaction. She currently works as a research assistant investigating bi/multilingual education, language learning, and autism in human-robot interaction

Nurziya Oralbayeva holds an M.A. degree in multilingual education from Nazarbayev University's Graduate School of Education. Currently, she is involved in research investigating the long-term effects of robot-assisted autism therapy.

Zhansaule Telisheva holds a B.S. in robotics and mechatronics from Nazarbayev University's School of Engineering and Digital Sciences. Her research interests include robot-assisted autism therapy and language learning in the field of human-robot interaction.

Aida Zhanatkyzy holds a B.S. in robotics and mechatronics from Nazarbayev University's School of Engineering and Digital Sciences. Her current research focuses on robot-assisted autism therapy and early literacy learning with robots and technologies.

Aidar Shakerimov received a B.S. degree in computer science and M.S. degree in robotics from Nazarbayev University in 2021 and 2023, respectively. He is currently working as a research assistant at the Institute of Smart Technologies and Artificial Intelligence (ISSAI) at Nazarbayev University. His research interest includes the application of reinforcement learning in human-robot interaction and robot control.

Shamil Sarmonov is a graduate of Nazarbayev University, holding a B.S. degree in robotics and mechatronics. He is currently pursuing his M.S. degree in robotics at Nazarbayev University. Additionally, Sarmonov serves as a research assistant at the Human-Robot Interaction (HRI) laboratory within Nazarbayev University. His research focuses on the application of learning algorithms in the fields of human-robot interaction and education.

Arna Aimysheva holds a B.S. in robotics and mechatronics from Nazarbayev University School of Engineering and Digital Sciences. Her current research interests include human-robot interaction, reciprocal peer tutoring, and language learning using social robots.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

1528-4972/23/09 $15.00