International Journal of Educational Technology and Learning
Vol. 19, No. 1, pp. 35-46, 2025
DOI: 10.55217/101.v18i3.961

Applying grey theory to evaluate the effectiveness of integrating hands-on robotics into english remedial instruction for fourth-grade underachievers

Ping-Huang Sheu

Department of Children English Education, National Taipei University of Education, Taiwan.
E-mail: samsheu@tea.ntue.edu.tw

Abstract

This study explores the effectiveness of humanoid robots as instructional tools in remedial English education for elementary school underachievers. The integration of educational robotics into language instruction has attracted growing interest, yet its application in supporting low-achieving learners remains underexplored. Nineteen fourth-grade students identified as underachievers participated in a robot-assisted learning program aimed at improving their English language performance. The study employed the S-P chart and GSP chart analytical methods to examine students' academic achievement and the relative difficulty of test items before and after the intervention. Analysis of the S and P curves revealed that the correct answer rate increased following instruction, with the test difficulty shifting to a moderate level. The caution index provided additional diagnostic insights into students' learning conditions and the performance consistency of specific test items. Furthermore, the GSP chart indicated a general improvement in student academic achievement, even though the test difficulty remained unchanged. These findings suggest that robot-assisted instruction can support learning gains among elementary underachievers and highlight the value of grey system theory in tracking both student progress and instructional effectiveness. The study offers practical implications for integrating robotics into language education and demonstrates the potential of grey system theory as a diagnostic tool in educational assessment, particularly in designing interventions for students requiring additional support.

Keywords:
English remedial instruction
GSP chart
Hands-on Robot
S-P chart.

Publisher:
Scientific Publishing Institute

Received: *15 June 2025*
Revised: *23 June 2025*
Accepted: *21 July 2025*
Published: *30 July 2025*

Funding: This study received no specific financial support.
Institutional Review Board Statement: The study involved minimal risk and followed standard ethical guidelines for social science fieldwork. Formal approval from an Institutional Review Board was not required due to the nature of the study. The study was done following basic ethical guidelines and principles. Informed verbal consent was obtained from all participants, and all data were anonymized to protect participant confidentiality.
Transparency: The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.

1. Introduction

The use of robotics as a pedagogical tool has become a prominent area of research in recent years (Klassner & Anderson, 2003; Ryu, Kwak, & KIM, 2008). Rooted in the theory of constructionism (Papert & Harel, 1991) this approach emphasizes learning through active design and assembly processes, allowing students to construct knowledge by engaging with tangible, programmable objects such as robots. Educational robots have been shown to be particularly effective in enhancing motivation and conceptual understanding, especially in mathematics and physics (Cooper, Keating, Harwin, & Dautenhahn, 1999). Beyond traditional science education, robotics is increasingly being integrated into a wide range of disciplines, including the arts and social sciences - an interdisciplinary movement often referred to as the robotic revolution (Hendler, 2000). This shift reflects the growing accessibility and versatility of robotic platforms, enabling even non-technical educators to incorporate them into their curricula.

Weinberg and Yu (2003) identified three key factors contributing to the educational effectiveness of robots: unique learning experiences, affordability, and ease of use. First, robots serve as physical embodiments of computational thinking, offering students an interactive environment in which to hypothesize, design, and test their ideas. This hands-on engagement fosters deeper cognitive and emotional connections to the learning material. For example, Turkle and Papert (1992) research on children's interactions with educational robots significantly influenced the development of learner-centered robotic tools.

Second, the declining cost of hardware has made robotics more accessible to schools with limited budgets. Affordable platforms such as the Handy Board and Lego Mindstorms RCX have simplified hardware interfacing and programming, thus lowering barriers to classroom integration (Marin, Mikhak, Resnick, Silverman, & Berg, 2000). Third, modern educational robots are designed for immediate usability with minimal setup. The term plug-and-play refers to devices that function immediately upon connection, without requiring complex technical configuration (Kanda, Hirano, Eaton, & Ishiguro, 2004). In short, the simplicity of setup enhances accessibility and facilitates the integration of robotics into educational settings. Taken together, these factors underscore robotics as a powerful and practical tool that offers multifaceted educational value rooted in experiential learning, cost-effectiveness, and user accessibility.

Numerous studies have examined the efficacy of instructional tools—such as multimedia and computer-assisted learning—in facilitating language acquisition, enhancing communication skills, and increasing learner motivation and engagement (Chinnery, 2006; Meskill & Anthony, 2005; Stockwell, 2007; Xie, Antle, & Motamedi, 2008; Yang & Chen, 2007). However, such tools often face limitations in adapting to learners' language levels, enabling direct interaction, and providing customization (Chang, Lee, Chao, Wang, & Chen, 2010; Lee, 2005; Wu, Chang, Liu, & Chen, 2008). Interactive robots have the potential to address these challenges through their embodied presence, multimodal interaction capabilities (e.g., speech, gestures, facial expressions), and adaptability to learner responses. For instance, humanoid robots can simulate peer-like or tutor-like roles in language learning environments, providing conversational practice, pronunciation feedback, and even emotional support (Kennedy, Baxter, & Belpaeme, 2016; Wollny, Schneider, & Kopp, 2021).

Despite these potentials, most applications of robotics in education remain focused on second language instruction at the secondary and tertiary levels, with few studies evaluating the use of robots in foreign language learning at the elementary level (Belpaeme, Kennedy, Ramachandran, Scassellati, & Tanaka, 2018; Chang et al., 2010). This gap presents both a challenge and an opportunity for future research to explore how robots can be effectively integrated into foreign language classrooms. Humanoid robots can execute students' programmed commands to move on a map and pronounce the English names of command cards through audio playback (Mubin, Stevens, Shahid, Al Mahmud, & Dong, 2013). This study leverages these two features by engaging students in designing map routes and recording self-introductions to enhance their English learning outcomes. Accordingly, the study investigates the use of humanoid robots as instructional tools for teaching English as a foreign language during remedial instruction at the elementary school level.

In addition, effective and appropriate assessment tools are essential for evaluating student learning progress and instructional effectiveness, as well as for providing crucial feedback on both teaching and learning processes (Greig, Taylor, & MacKay, 2012). While statistical analysis is commonly regarded as a valuable tool, it requires sufficiently large datasets to minimize bias and ensure accuracy (Barnett et al., 2018). However, because statistical results reflect trends at the group level rather than individual performance, they are not ideally suited for remedial instruction, which emphasizes the progress of individual underachieving students (Jones & Coffey, 2016).

Sato (1969) introduced the Student-Problem (S-P) chart as a graphical tool to help teachers diagnose students' learning achievements and difficulties by illustrating the relationship between students and specific learning problems. This provides diagnostic data that enables teachers to adapt their instructional strategies accordingly (Sheu, Pham, Nguyen, & Nguyen, 2013). According to Grey System Theory (Deng, 1982) the Grey Student-Problem (GSP) chart can further identify and clarify uncertain and influential factors in student performance before, during, and after instruction. Accordingly, this study adopts both the S-P chart and GSP chart to diagnose and analyse the effects of remedial instruction on fourth-grade students' English learning performance.

The purpose of this study was to investigate the impact of integrating the T. Robot into English remedial instruction on fourth graders' English achievement and test quality. It employed the S-P chart to visualize the position of each student and question item, and the GSP chart to reveal student performance and item difficulty. The research questions were as follows:

1. Are there differences in the S-P charts before and after remedial instruction?

2. To what extent do the types of students' learning profiles and difficulty of question items change after remedial instruction?
3. To what extent does the relationship between students' performance and test item difficulty change after remedial instruction?

2. Method

2.1. Participants and Instrument

English education has been incorporated into six grades in elementary schools in Taiwan, and students learn English as a foreign language with three 40-minute period of English class per week via textbooks to cultivate basal listening and speaking skills, with gradually move on reading and writing. Mid-term and final test are used for assessing their achievement, and based on the test results, those who are under average scores are required to attend an after-school remedial program at least for one semester.

Nineteen students were selected from five fourth-grade classes at an elementary school in Taipei City. These students had received formal English education in school for at least three years. Mandarin was their primary language, while they spoke regional dialects such as Taigi or Hakka at home. The participants attended one 40-minute remedial session every Wednesday afternoon for at least one semester.

The English test in this study was developed based on the fourth-grade English textbook used in the school curriculum and consisted of 12 items in total. The test measured participants' improvement in vocabulary and basic sentence structures, and included tasks such as filling in blanks using contextual cues from texts and images. The same test version, with a different item order to minimize memory effects, was administered before and after the robot-assisted intervention. The test demonstrated high reliability, with Cronbach's alpha coefficients of 0.91 for the pre-test and 0.90 for the post-test.

2.2. Robot-Assisted Remedial Program

The remedial program served as a supplementary intervention, consisting of 40-minute weekly lessons over a period of seven weeks. The lessons were designed based on the content of the school's fourth-grade English textbook.

2.2.1. T. Robot Description

T. Robot (model number AP-708) is an intelligent educational robot developed in 2018 through a collaboration among the Institute for Information Industry (III), Reading & Rhythm, AmPo International, and Sonix Technology. It was specifically designed as an introductory programming tool for children aged 5 to 8, featuring a compact humanoid form approximately the size of an adult's palm. The robot does not have movable arms. The front side resembles a human face and torso, while the back side includes a speaker outlet, an OID optical recognition sensor, a power switch, and a charging port arranged from top to bottom (as shown in the left panel of Figure 1). An LED light is located on the top of the robot, and a set of wheels on the bottom enables mobility across a mapped surface (as shown in the right panel of Figure 1).

Figure 1. The appearance and map of Billard and Kragic (2019).

T. Robot employs optical recognition technology to read programming command cards and is compatible with 79 different types of cards (see Table 1), including movement cards, equipment cards, and sound effect cards (see Figure 2). These cards allow the robot to execute a variety of commands such as movement, lighting, sound effects, and playing pre-recorded audio files. Each card includes both an image and corresponding English text to support language development. When a card is read, the robot audibly pronounces the name of the card. Once activated, the robot responds exclusively to the sequence of command cards arranged by the user. Students are first taught to correctly sequence the cards to control the robot's actions.

Table 1. Types of command cards and their functions.

Type	Colour	Number	Functions
Basic cards	Red	16	initiate and terminate the program, adjusting volume, turn the light on or off, changing light colours, and record audio
Movement cards	Fellow	12	the number of steps and directions for robot movement
Equipment cards	Blue	18	play songs, sett timers (In seconds), stop actions, and operate gripping mechanisms
Sound effect cards	Green	23	produce animal sounds and transportation-related sound effects
Advanced cards	Orange	10	program commands such as loops, save and delete programs, and program checking functions

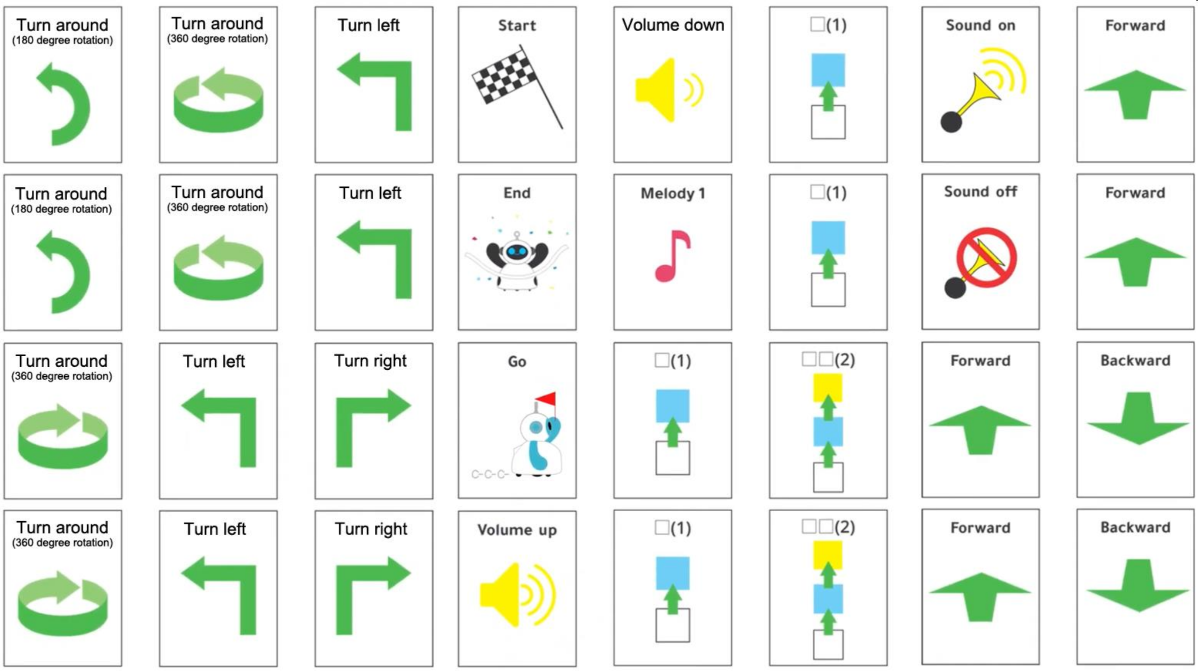

Figure 2. Command cards.

After sequencing the command cards, Billard and Kragic (2019) is placed on the map to perform the programmed instructions. The large-format version of the basic map consists of a grid system and can be used in conjunction with worksheets or physical objects. Students may mark routes or assign object locations on the worksheets, or construct three-dimensional objects directly on the map. These features enhance task clarity and help students better understand the robot's missions.

2.2.2. Robot-Assisted Instruction

The 7-week T. Robot instructional program was task-based and centered on Units 1 ("*A Surprise for Jello*") and 2 ("*City Adventure*") from *Wonder World*, published by KST Education Corp. The instructional sequence is described as follows:

Week 1: The first session introduced the course, providing an overview of the course theme, schedule, and group activity format. Students also reviewed how to operate the T. Robot, practicing basic directional and locational commands (i.e., *up*, *down*, *left*, and *right*) as well as step-count instructions.

Week 2: Following instruction on Unit 1 ("*A Surprise for Jello*"), students used worksheets to mark the positions of various places and label them with corresponding English vocabulary such as *bank*, *bookstore*, *hospital*, *park*, *post office*, *restaurant*, and *supermarket*. They then composed route phrases using recording cards, enabling the robot to represent the street layout accurately. Through collaborative discussion, students worked in groups to complete the task and documented their work by photographing their arrangements and uploading them to the Seesaw application.

Week 3: Students wrote command card programs that enabled the robot to navigate to each place and introduce its name. The focus then shifted to programming the robot to introduce these locations and using recording cards to capture the corresponding spoken output. For example, the robot might say, "*This is the post office. People can mail letters to their friends.*"

Week 4: Using the same route planned previously, students had the robot introduce each place and its function. Groups then discussed and determined the relative positions of other places and used recording cards to capture related descriptions. Students subsequently wrote programs and recorded audio that enabled the robot to introduce the places and describe their functions. These outputs were recorded and uploaded to the Seesaw platform.

Week 5: After introducing Unit 2 ("*City Adventure*"), students listed the places from the previous unit and added modes of transportation (*bike*, *bus*, *car*, *MRT*, *plane*, *scooter*, *taxi*, and *train*), followed by route phrases (e.g., "*go to the hospital by taxi*", "*go to the park by bike*"). These phrases were recorded using the robot's recording function, arranged chronologically, and mapped onto specific locations on a grid-based map. Using worksheets, students designed the layout of their routes on the map. This activity involved two robots working in relay to manage the increased complexity and time demands. Each group used two sets of command cards corresponding to the respective robot sequences and began programming the route design for a dual-robot presentation.

Week 6: Students programmed a robot to navigate the route map they had designed. Upon reaching each designated location, the robot announced the corresponding route. Students then used movement, recording, and sound cards to create complete sentences combining places with modes of transportation (e.g., "*I go to the bookstore by bus*"). Instruction cards enabled the robot to move between locations, announce the places, and play the recorded audio.

Week 7: Students first reviewed and practiced the tasks from the previous weeks. They then used two sets of cards to program and operate two robots, each presenting the route on the map, including both places and modes of transportation.

2.3. Data Analysis

2.3.1. Grey System Theory (GST)

Deng (1982) proposed grey system theory for solving uncertain problems with partially known information, which can be classified into three categories: white, grey and black systems. It is identified as an effective methodology that can be used to solve uncertain problems through model construction, grey prediction, grey relational analysis, decision making and grey control (Nagai, Yamaguchi, & Li, 2005).

2.3.2. Student-Problem (S-P) Chart

Sato (1969) utilize calculated and sorted coefficient to construct a Student-Problem (S-P) chart, in which caution index (CI) for students and for item (CP) were used to decide whether individual student's or question's response patterns are anomalies (i.e. unusual or aberrant) (Nagai et al., 2005). S-P Chart can not only display the diagnostic assessment of student learning, but also present the effectiveness of instructions (McArthur, 1983). It diagnoses, analyzes, processes and arranges data in a defined order (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013; Yu & Yu, 2006) so as to identify the quality of test items and diagnosis of students' learning. The constructing process of a S-P chart is as follows (Nguyen et al., 2013).

Step 1: Mark correct answer as 1 and wrong one 0 to construct a matrix structure of student-problem.

Step 2: Calculate caution index for students (*CS*) and caution index for items (*CP*).

Step 3: Arrange the caution index for students (*CS*) from high to low on the vertical column and caution index for items (*CP*) from more to less on the horizontal row above, to from the S-P chart as shown in Figure 2 below.

Student number	Problem number P_j, $j=1,2,\cdots,n$	Total score	CS
S_i, $i=1,2,\cdots,m$	$Y=(y_{ij})$	High ↕ SS_i Low	CS_i
Number of correct answer	More ↔ Less PP_j	$\sum_{i=1}^{m} SS_i = \sum_{j=1}^{n} PP_j$	–
CP	CP_j	–	–

Fiqure 2. S-P chart.

Step 4: Plot each student's caution index and the percentage of scores on the learning type diagram (Left in Figure 3) to determine the type of student learning, and plot problem caution index and the percentage of correct responses on the problem type diagram (Right in Figure 3) to diagnose the type of question item (Sheu et al., 2013).

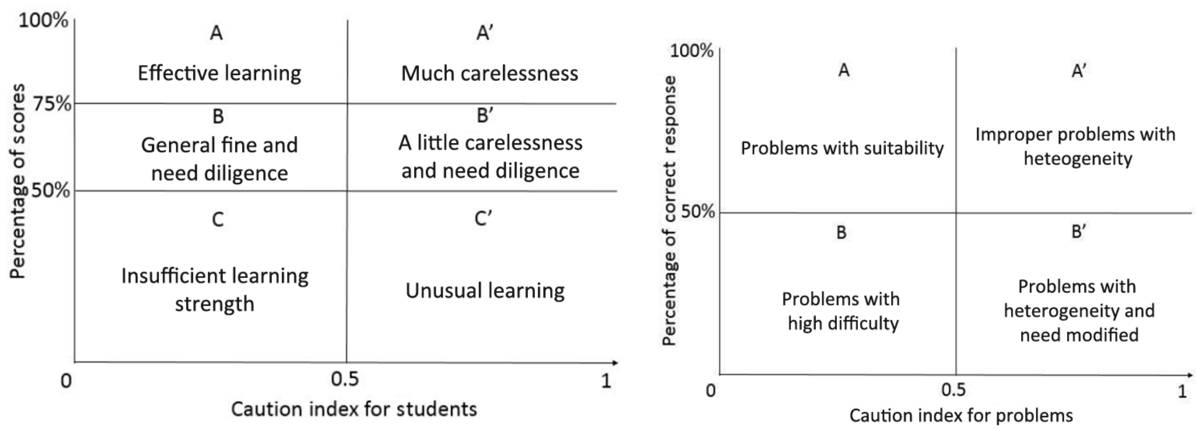

Figure 3. Student (left) and question (right) diagnostic analysis.

2.3.3. Grey Student-Problem (GSP) Chart

The GSP Chart is a visual and analytical tool used to assess the relationship between students and problems (or tasks/exams/subjects). It evaluates how well a student performs on different problems, helping educators identify strengths and weaknesses in both students and test items. Nagai et al. (2005) combine S-P chart and GRA to establishes a grey student-problem (GSP) Chart so as to deal with the insufficiency and uncertain factors of the S-P chart. The Algorithm of GSP chart is as follow (Nguyen et al., 2013):

Step 1: Plot students' LGRA-S (columns) and LGRA-P (rows) to construct the GSP binary matrix as shown in the following Figure 4 .

Student number	Problem number P_j , $j = 1,2,\cdots,n$	Total score	*LGRA-S*
S_i , $i = 1,2,\cdots,m$	$X = [x_{ij}]_{m \times n}$	SS_i	High ↕ GS_i Low
Number of correct answer	PP_j	$\sum_{i=1}^{m} SS_i = \sum_{j=1}^{n} PP_j$	–
LGRA-P	More ↔ Less GP_j	–	–

Figure 4. GSP chart matrix.

Step 2: Calculate each student-problem pair's grey relational coefficient (*GRC*), and then generate the local grey relational grade (*GRG*) of the i-th student (*GSi*) and of the j-th problem (*GPj*).

Step 3: Aggregate the coefficients for each row (student) or column (problem) to determine overall performance or difficulty. Sort *GSi* value ($i = 1, 2\ldots m$) from high to low and students' GRG from strong to weak, and *GPj* value ($j = 1, 2\ldots n$) from low to high and problems' GRG from easy to hard.

Step 4: Draw the GSP Chart to demonstrate each student's performance and item difficulty and localized grey relational grade to establish the GSP chart.

3. Results

3.1. S Curve and P Curve Analysis

The S curve (blue line) and P curve (red line) present the average percentage of correct answers in the pre-test and post-test (shown in Table 2). As can be seen, areas on the left of both curves show that the correct answers rate is spread from upper-left (about 35%) to central right (about 70%). This result indicates that this test has not very difficult questions neither very easy questions after instruction.

Table 2. S curve and P curve of the pre-test and post-test.

Pre-test												Post-test											
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	0	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1
1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1
1	1	1	1	1	0	1	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	0
1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0
1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	1
1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	0	1	0	1	1
0	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	1	0	0
1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	1	0	0	0
1	1	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	0	1	0	0
1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	0	1	0	0	0	1
1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	1	0	0	0	1	0
0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

3.2. Student Chart Analysis

Table 3 presents the general information (i.e. total number and percentage of correct answers, and caution index) of student's S-P chart. In the pre-test, S_6 got the highest score while five students (S_{10}, S_{15}, S_{16}, S_{12}, S_{21}) received the lowest one; in the post-test, three students (S_5, S_{11}, S_{20}) got the highest score whereas S_{12} remained at the same position. Regarding the caution index (CS), none was greater than 0.5 in the pre-test, but in the post-test, three students (S_{19}, S_{13}, S_1) achieved this level, indicating that they needed teacher's attention in order to avoid carelessness in learning.

Table 3. Students assessment results of S-P chart.

Pre-test	Total	Ratio	CS	Type	Post-test	Total	Ratio	CS	Type
30306	12	1.00	0.00	A	30305	12	1.00	0.00	A
30320	11	0.92	0.20	A	30311	12	1.00	0.00	A
30305	9	0.75	0.15	A	30320	12	1.00	0.00	A
30319	8	0.67	0.17	B	30306	11	0.92	0.00	A
30311	6	0.50	0.00	B	30319	11	0.92	1.29*	A'
30318	6	0.50	0.00	B	30303	10	0.83	0.64*	A'
30303	4	0.33	0.00	C	30304	10	0.83	0.00	A
30313	4	0.33	0.16	C	30313	10	0.83	0.00	A
30314	4	0.33	0.00	C	30301	9	0.75	0.64*	A'
30301	3	0.25	0.30	C	30314	8	0.67	0.32	B
30304	3	0.25	0.00	C	30315	8	0.67	0.06	B
30309	3	0.25	0.30	C	30318	8	0.67	0.32	B
30317	3	0.25	0.00	C	30307	7	0.58	0.31	B
30307	2	0.17	0.00	C	30321	7	0.58	0.31	B
30310	1	0.08	0.33	C	30309	6	0.50	0.06	B
30312	1	0.08	0.00	C	30310	4	0.33	0.31	C
30315	1	0.08	0.00	C	30317	4	0.33	0.00	C
30316	1	0.08	0.00	C	30316	2	0.17	0.00	C
30321	1	0.08	0.11	C	30312	1	0.08	0.00	C

Note: * CS > 0.5.

3.3. Problem Chart Analysis

Table 4 shows the general information (i.e. total number and percentage of correct answers, and caution index) of question's S-P chart. When the level of question items is considered, in the pre-test, P_6 was the easiest question, and P_{12} and P_4 were the most difficult question. However, in the post-test, P_{12} was the easiest question, but P_{10}, P_6 and P_{11} were the hardest one. When the caution index (CP) is considered, none was greater than 0.5 in both tests, suggesting that both tests were appropriate to measure students' achievement.

Table 4. Question assessment results of S-P chart.

Pre-test	Total	Ratio	CS	Type	Post-test	Total	Ratio	CS	Type
6	16	0.84	0.20	A	12	19	1	0.00	A
2	15	0.79	0.00	A	5	18	0.95	0.00	A
3	13	0.68	0.10	A	7	16	0.84	0.00	A
1	9	0.47	0.04	B	2	14	0.74	0.48	A
10	6	0.32	0.19	B	3	13	0.68	0.04	A
5	5	0.26	0.21	B	9	13	0.68	0.25	A
11	5	0.26	0.00	B	1	12	0.63	0.20	A
9	4	0.21	0.00	B	8	12	0.63	0.16	A
7	3	0.16	0.00	B	10	11	0.58	0.00	A
8	3	0.16	0.16	B	4	8	0.42	0.21	B
4	2	0.11	0.00	B	6	8	0.42	0.17	B
12	2	0.11	0.00	B	11	8	0.42	0.17	B

3.4. Types of Students' Learning

The results of students' learning type calculated by caution index for students (CS) and percentage of scores are shown in Figure 5. In the post-test, three students (S_6, S_{20}, S_5) were categorized into Type A (effective learning), but in the post-test, the number of Type A was increased from 3 to 5, indicating that their learning was stable. Among them, one (S_{11}) was ascended from Type B (general fine and need diligence), and two (S_4, S_{13}) were from Type C (insufficient learning strength). It should be noted that two students (S_3, S_1) in

Type C were Typed A' (much carelessness), meaning that their learning was slightly uneven due to the lack of attention in learning. Also, one student (S_{19}) cannot be categorized into any of the six types in the post-test. After instruction, the number of Type B (general fine and need diligence) was increased from 3 to 6, implying that they need to try harder in learning. In the pre-test, Type B has three students (S_{18}, S_{11}, S_{19}), but one student (S_{18}) remained in the same condition, and the other five (S_{14}, S_{15}, S_{7}, S_{21}, S_{9}) from Type C (insufficient learning strength) were found in this type. Regarding Type C (insufficient learning strength), four students (S_{12}, S_{16}, S_{17}, S_{10}) stayed at the same condition in the pre-test and post-test, revealing that they learning is inadequate and lack of academic ability.

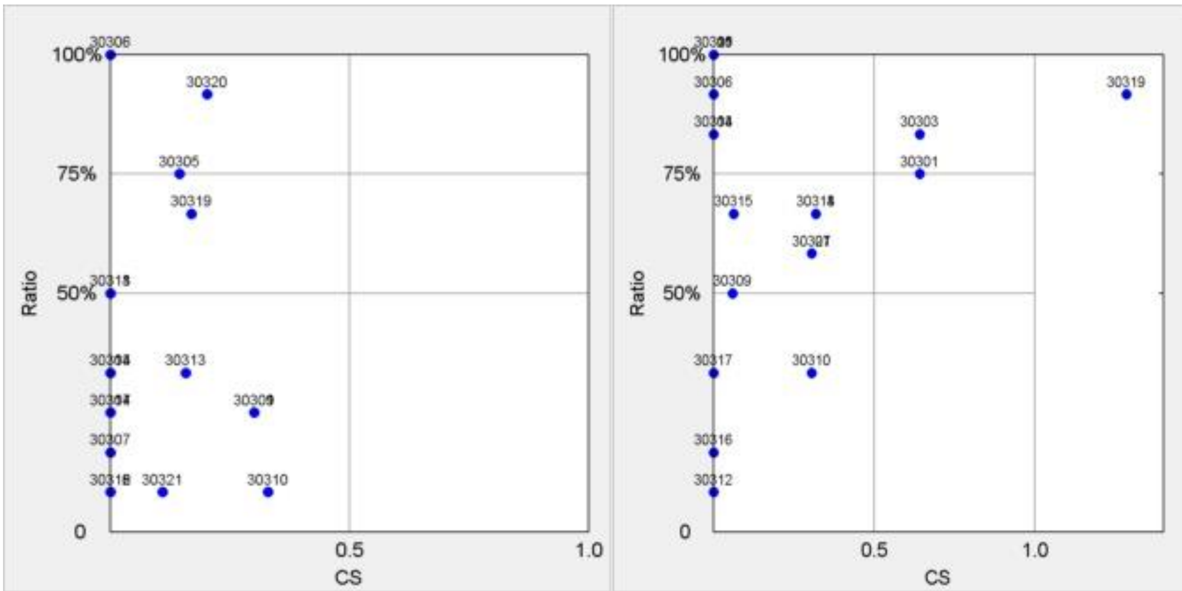

Figure 5. Student type diagram for pre-test (Left) and post-test (Right).

3.5. Types of Question

Caution index for problem (CP) and problem passing ratio were analyzed to diagnose the question quality of pre-test and post-test, and question type diagrams are shown in Figure 6. There were three questions (P_6, P_2, P_3) in Type A (problems with Suitability) in the pre-test; however, the number showed a tremendous increase of 9 questions in the post-test, signifying that these questions were suitable for distinguishing low achievers with different types of learners. Among them, one (P_6) is fallen into Type B (problems with high difficulty), and seven (P_{12}, P_5, P_7, P_9, P_1, P_8, P_{10}) from Type B were ascended into Type A. Moreover, in the pre-test, nine questions (P_1, P_{10}, P_5, P_{11}, P_9, P_7, P_8, P_4, P_{12}) typed B, but two questions (P_4, P_{11}) still stayed at the same difficult level, meaning they were very difficult and can be used for distinguishing high achievers.

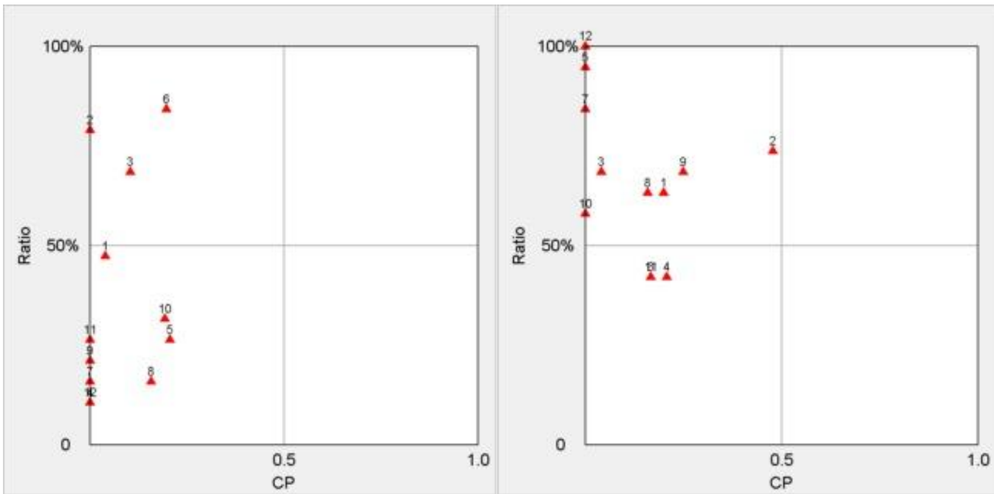

Figure 6. Question type diagram for pre-test (left) and post-test (right).

3.6. GSP Graph Analysis

The student achievement and question difficulty of pre-test and post-test are evaluated and classified. GSP graph of both tests are presented in Figure 7 and accordingly, the classification results of achievement and difficulty are summarized in Table 5. As can be seen, student achievements were more at bottom inferior level in the pre-test, whereas their performances scattered equally at the five levels in the post-test. This demonstrates that student academic achievement after instruction was very good and much better than the pre-test. Since the same test was used as pre- and post-test, the classification results of both tests were not so different. However, when the difficult level of both tests was similar and students performed better in the post-test, this signifies the improvement of students' learning and ability, and the positive effect of instruction. Although the difficulty of tests is the same, the questions in the post-test became easier than those in the pre-test.

Table 5. The classification results of students and questions.

	GS			GP		
	Achievement	**Pre-test**	**Post-test**	**Difficulty**	**Pre-test**	**Post-test**
0.8 - 1.0	A (Excellent)	1	3	A (Very easy)	2	1
0.6 - 0.79	B (Good)	1	5	B (Easy)	1	1
0.4 - 0.59	C (Medium)	2	4	C (Moderate)	0	1
0.2 - 0.39	D (Weak)	2	3	D (Difficult)	2	3
0.0 - 0.19	E (Inferior)	15	6	E (Very difficult)	7	6

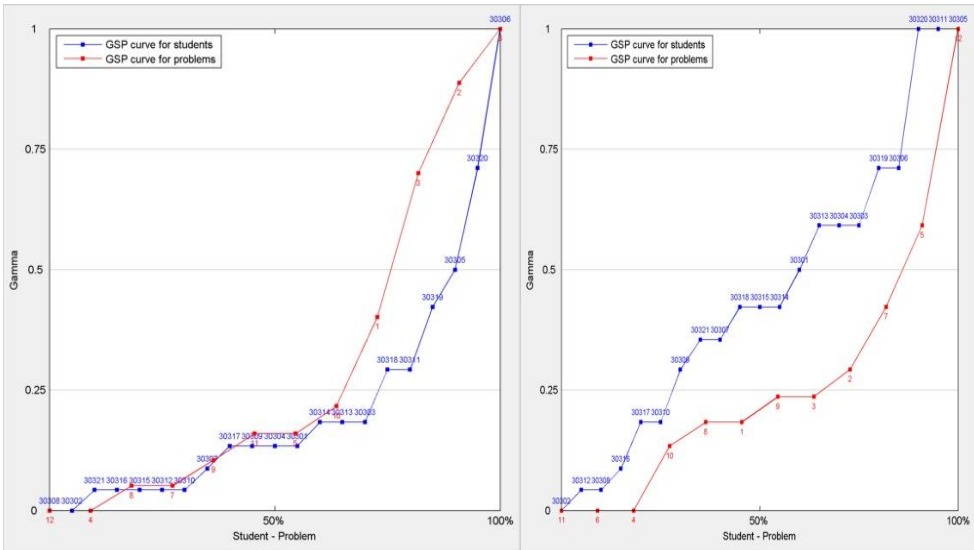

Figure 7. GSP Graph of pre-test and post-test.

4. Discussion and Implication

In this study, the S-P chart and GSP chart were employed to analyze and diagnose the effects of robot-assisted remedial instruction on fourth-grade underachievers in Taiwan. The primary aim was to illustrate improvements in individual students' learning performance and to assess item difficulty through the interpretation of these visual analytical tools.

Firstly, the S-P chart depicted students' accuracy rates alongside test item difficulty. The S-curve and P-curve diagonally segmented the chart, indicating areas of low (left) and high (right) achievement and difficulty, respectively. In this research, the observed shift of the data cluster from the upper-left to the middle region suggested an increase in students' correct response rates, implying that the test was moderately challenging and suitable for evaluating student learning. As the effectiveness of instruction is often obscured when a test is overly difficult or easy, the S-P chart proved valuable in illustrating this relationship.

Secondly, the Caution Index for Students (CS) was utilized to identify and classify students' learning profiles. The categorization into six learner types facilitated a clearer understanding of students' post-instruction conditions and enabled teachers to make adjustments to their teaching strategies. For instance, students who remained in Type B (generally fine but require diligence) or Type C (insufficient learning effort) after instruction warrant closer attention. It is essential for educators to adopt alternative instructional materials and activities to better engage these students and enhance their motivation.

Thirdly, the Caution Index for Problems (CP) was applied to evaluate the quality of test items, categorizing them into four distinct types. This classification allowed educators to discern the characteristics of each item, particularly in distinguishing between high- and low-achieving students. Students who answered

Type B items (i.e., those with high difficulty) correctly on the post-test generally performed better overall, suggesting that such items can serve as effective discriminators of achievement levels.

Furthermore, the GSP charts provided a visual representation of individual student progress by comparing pre- and post-test results. These charts highlighted the relationship between student performance and item difficulty, with the slope of the lines and corresponding Gamma values indicating improvement in post-test scores, even though item difficulty remained constant. In contrast to the S-P chart, the GSP chart offered deeper insight into students' developmental trajectories in response to instruction.

One implication of this research is that by designing, programming, and engaging in robot-assisted tasks, underachieving students became more actively involved in their learning. This active participation fostered improvements in both their language skills and learning behaviors. Consequently, elementary school teachers are encouraged to incorporate such instructional methods, and policymakers should consider exploring and providing diverse robotic tools for classroom use.

Another implication concerns the lack of standardized assessment tools for diagnosing and evaluating elementary school students' performance in Taiwan. This study highlights the utility of the S-P and GSP chart methodologies as practical and informative tools for formative assessment. Teachers are encouraged to adopt these methods to enhance instructional effectiveness and support student learning outcomes.

In conclusion, the robot used in this study functioned as an instructional assistant that effectively engaged underachieving students in learning activities. The findings underscore the potential of robotics to address challenges in English language learning, promote collaboration and problem-solving, and increase classroom engagement. The combined use of S-P and GSP charts served as a valuable assessment framework, providing visual data to inform instructional adjustments. This research offers meaningful insights for elementary English teachers and remedial instructors, particularly in evaluating the effectiveness of interventions and assessing student progress in English language acquisition.

References

Barnett, E. A., Bergman, P., Kopko, E., Reddy, V., Belfield, C. R., & Roy, S. (2018). *Multiple measures placement using data analytics: An implementation and early impacts report*. New York: Center for the Analysis of Postsecondary Readiness.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science robotics, 3*(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954

Billard, A., & Kragic, D. (2019). Trends and challenges in robot manipulation. *Science, 364*(6446), eaat8414.

Chang, C.-W., Lee, J.-H., Chao, P.-Y., Wang, C.-Y., & Chen, G.-D. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Journal of Educational Technology & Society, 13*(2), 13-24.

Chinnery, G. (2006). Emerging technologies - Going to the MALL: Mobile assisted language learning. *Language Learning & Technology, 10*(1), 9–16.

Cooper, M., Keating, D., Harwin, W., & Dautenhahn, K. (1999). Robots in the classroom: Tools for accessible education. In Buhler, C. & Knops, H. (Eds.), Assistive technology on the threshold of the new millennium. In (pp. 448–452). Amsterdam: IOS Press.

Deng, J. L. (1982). Control problems of gray system. *Systems &Control Letters, 1*(5), 288-294. https://doi.org/10.1016/S0167-6911(82)80025-X

Greig, A., Taylor, D. J., & MacKay, T. (2012). *Doing research with children: A practical guide*. London, UK: Sage.

Hendler, J. (2000). Robots for the rest of us: Designing systems out of the box. In A. Druin & J. Hendler (Eds.), Robots for kids: Exploring new technologies for learning. In (pp. 2-7). San Mateo, CA: Morgan Kaufmann.

Jones, J., & Coffey, S. (2016). *Modern foreign languages 5–11: A guide for teachers*. London, UK: Taylor and Francis.

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human–Computer Interaction, 19*(1-2), 61-84.

Kennedy, J., Baxter, P., & Belpaeme, T. (2016). *The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning*. Paper presented at the In Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (pp. 67–74).

Klassner, F., & Anderson, S. D. (2003). LEGO mindStorms: Not just for K–12 anymore. *IEEE Robotics & Automation Magazine, 10*(2), 12–18. https://doi.org/10.1109/MRA.2003.1213611

Lee, S. M. (2005). The pros and cons of task-based instruction in elementary English classes. *English Teaching, 60*(2), 185–205.

Marin, F., Mikhak, B., Resnick, M., Silverman, B., & Berg, R. (2000). To mindstorms and beyond: Evolution of a construction kit for magical machines. In A. Druin & J. Hendler (Eds.), Robots for kids: Exploring new technologies for learning. In (pp. 9–33). San Mateo, CA: Morgan Kaufmann.

McArthur, T. (1983). *A foundation course for language teachers*. Cambridge: Cambridge University Press.

Meskill, C., & Anthony, N. (2005). Foreign language learning with CMC: Forms of online instructional discourse in a hybrid Russian class. *System, 33*(1), 89-105.

Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J.-J. (2013). A review of the applicability of robots in education. *Journal of Technology in Education and Learning, 1*(209-0015), 13. https://doi.org/10.2316/Journal.209.2013.1.209-0015

Nagai, M., Yamaguchi, D., & Li, G.-D. (2005). Grey structural modeling. *Journal of Grey System, 8*(2), 119-130.

Nguyen, P.-T., Nguyen, P.-H., Pham, D.-H., Tsai, C.-P., & Nagai, M. (2013). The proposal for application of several grey methods in evaluating and improving the academic achievement of students. *International Journal of Kansei Information, 4*(4), 179-189.

Papert, S., & Harel, I. (1991). Situating constructionism. In I. Harel & S. Papert (Eds.), Constructionism. In (pp. 1–11). Norwood, NJ: Ablex Publishing.

Ryu, H. J., Kwak, S. S., & KIM, M. S. (2008). Design factors for external form of robots as elementary school teaching assistants. *Bulletin of Japanese Society for the Science of Design, 54*(6), 39-48.

Sato, T. (1969). *A method of analyzing data gathered by the response analyzer for diagnosis of student performance and the quality of instructional sequence.* Paper presented at the Proceedings of IECE of Japan annual conference, S12-1.

Sheu, T.-W., Pham, D.-H., Nguyen, P.-T., & Nguyen, P.-H. (2013). A matlab toolbox for student-problem chart and grey student-problem chart and its application. *International Journal of Kansei Information, 4*(2), 75-85.

Stockwell, G. (2007). A review of technology choice for teaching language skills and areas in the CALL literature. *ReCALL, 19*(2), 105-120. https://doi.org/10.1017/S0958344007000225

Turkle, S., & Papert, S. (1992). Epistemological pluralism and the revaluation of the concrete. *Journal of Mathematical Behavior, 11*(1), 3-33.

Weinberg, J. B., & Yu, X. (2003). Robotics in education: Low-cost platforms for teaching integrated systems. *IEEE Robotics & Automation Magazine, 10*(2), 4-6. https://doi.org/10.1109/MRA.2003.1213610

Wollny, S., Schneider, S., & Kopp, S. (2021). Embodied tutors for language learning: The role of social cues and feedback strategies. *Frontiers in Robotics and AI, 8*, 626519.

Wu, C. C., Chang, C. W., Liu, B. J., & Chen, G. D. (2008). *Improving vocabulary acquisition by designing a storytelling robot.* Paper presented at the In Proceedings of 8th IEEE International Conference on Advanced Learning Technologies (pp. 498–500). Washington, DC: IEEE Computer Society.

Xie, L., Antle, A. N., & Motamedi, N. (2008). *Are tangibles more fun? Comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces.* Paper presented at the In TEI '08: Proceedings of the 2nd International Conference on Tangible and Embedded Interaction (pp. 191–198). New York: ACM.

Yang, S. C., & Chen, Y.-J. (2007). Technology-enhanced language learning: A case study. *Computers in Human Behavior, 23*(1), 860-879. https://doi.org/10.1016/j.chb.2006.02.015

Yu, S., & Yu, M. (2006). The relationships among indices of diagnostic assessment of knowledge Structure and SP chart analysis. *Journal of Education & Psychology, 29*(1), 183-208.