

# Applying Student-Problem Chart, Grey Student-Problem Chart and Grey Structure Modeling to Analyze the Effect of an Elementary School English Remedial Instruction

Ping Huang Sheu<sup>1</sup>

<sup>1</sup> Department of Children English Education, National Taipei University of Education, Taipei, Taiwan

Correspondence: Ping Huang Sheu, 134, Sec 2, Heping East Rd, Daan District, Taipei 10671, Taiwan. E-mail: samsheu@tea.ntue.edu.tw

Received: August 23, 2019      Accepted: September 20, 2019      Online Published: October 13, 2019

doi:10.5539/ijel.v9n6p49      URL: <https://doi.org/10.5539/ijel.v9n6p49>

## Abstract

Student-Problem (S-P) chart, Grey Student-Problem (GSP) chart and Grey Structural Modeling (GSM) are graphical analysis tools to represent the relationship between students and test items which can be applied to diagnose, analyze, and evaluate students' learning situation and achievement. This study has adopted a combination of these graphic analyses to assess the effect of an elementary school remedial instruction. A group of 20 fourth graders in Taipei city participated in this study, and an English test, consisting of four sections with 25 question items in total, which is based on the textbook in the fourth grade, was conducted before and after the remedial instruction. The results indicated that S-P chart, the distribution diagram of student and item types, GSP and GSM graphs effectively diagnosed students' learning condition and items' difficulty, and demonstrated the effect of students' achievement and the change of test items' types. It is strongly recommended that teachers utilize these methods in assessing the progress of their teaching and their students' learning.

**Keywords:** S-P chart, GSP chart, GSM graph, English remedial instruction

## 1. Introduction

Since English has been included as a school subject in elementary schools in Taiwan for at least 15 years, so-called two-peak phenomena of students' English proficiency has become obvious, that is, high achievers get better and better and low achievers become worse and worse. In Grade 9 curriculum, Ministry of Education (2003) states expressly that teachers should implement diagnostic assessment and remedial instruction in order to help those who cannot reach average level of English achievement in the school. The focus is to help these underachievers learn what they were not able to learn in English lessons, yet teachers might adopt different activities to do so. This has rendered diagnostic instruments vital and essential for teachers so that they can identify their students' English learning problems and accordingly, design appropriate English remedial instructions (Sheu, 2019).

Generally speaking, most elementary school English teachers use mid- and final-term tests as a tool for assessing students' achievement. By doing this, they can know if their students have learned the content of textbooks; however, it is far too simple to draw conclusions from this about students' achievement (Sheu, 2017). This is because English teachers construct mid- and final-term tests on the basis of their own experiences and the suitability and reliability of tests are not normally considered. In this case, students' actual learning progress and difficulties would not be possibly identified (Salimi & Farsi, 2018). If teachers therefore design remedial instructions, the content, methods and activities are highly inadequate, and the effects are definitely questionable (Kirsch, 2008; Mitchell & Myles, 2004).

Regarding the assessment of English remedial instructions in Taiwan, there is not an official format or measurement of assessment for teachers to be followed. It is decided by each individual school and mostly, the instructor makes the decision what to be assessed and how it is assessed, and thus, the effect of these instructions is varied and uncertain (Sheu, 2015). In fact, there is no a single test as the assessment instrument of remedial instructions for all schools when the nature of students, schools and locations is considered (Csizér & Magid, 2015; Jacobs & Farrell, 2001). Putting it another way, it is more practicable to have an analytic method for identifying the efficiency of remedial instructions rather than to have an individual test for assessing the effect of

all instructions.

Adding these together, without an appropriate assessment tool, students' learning progress and difficulties would be indeterminable; if there is a lack of an effective analytic method, conclusions about the effect of any remedial instruction is unrealistic (Greig, Taylor, & MacKay, 2012). Thus, having such a method appears to be pivotal in the whole remedial instruction and consequently, the implementation of such an instruction will have profound effects on the entire English learning and teaching process (Sheu, 2016).

Statistics analysis is commonly conducted as an assessment instrument to evaluate the effect of English language teaching and learning, but it requires a certain number of samples and a period of time in order to avoid bias and errors in the data and results (Barnett, Bergman, Kopko, Reddy, Belfield, & Roy, 2018). In other words, credible instrument and results rest heavily upon proper design and implementation. The teaching loading in elementary schools and the time-consuming nature of analyzing data make it difficult for teachers to do teaching and analysis at the same time (Sheu, 2017). What is more, since the statistic results have to take into account the whole rather than an individual student, they are not relevant to the intention of remedial instruction to concentrate on the progress of each underachiever. In this sense, analytic methods play a decisive role in implementing and evaluating remedial instructions appropriately and effectively (Jones & Coffey, 2016). In other words, analytic results provide English language teachers with crucial feedback on the process of their teaching and students' learning.

Sato (1969) used Student-Problem (S-P) chart to help teachers diagnose students' problems in their learning performance. As a graphical analysis, the chart revealed effectively the relationship between students and their learning problems. The main purpose of this chart is to get the diagnostic data of each student so that teachers can make any changes necessary to their teaching and students' learning (Sheu, Pham, Nguyen, & Nguyen, 2013). Based on Grey system theory (Deng, 1989), Tsai and Chung (2003) utilizes Grey Student-Problem (GSP) chart to clarify uncertain factors which affect teaching practice and to solve problems students encountered in their learning situation. That is, the GSP chart is applied to draw graphs to examine and compare students' performance before, during and after a learning situation. Grey Structure Modeling (GSM) is designed for clustering students into different groups based on their performance in a test (Yamaguchi, Li, & Nagai, 2005). It helps teachers understand their students' learning process and achievement better (Wang, Sheu, Liang, Tzeng, & Nagai, 2011). Previous research has demonstrated the effect of these graphs in various learning and teaching programs (Chen, Lai, & Liu, 2005; Chen, Lai, & Jhan, 2007; Lin & Chen, 2006; Sheu, Tzeng, Liang, Wang, & Nagai, 2012). Thus, this paper adopts S-P chart, GSP chart and GSM graph to diagnose and analyze the effect of a remedial instruction on 4<sup>th</sup> graders' English learning.

### 1.1 Problem Student-Problem (S-P) Chart

Sato (1969) carried out analysis and sorting based on students' response pattern in test items, and calculated coefficient to construct a Student-Problem (S-P) chart, in which caution index for students (CS) and for item (CP) were used to judge whether individual students or questions are anomalies in response patterns (Sheu, Chen, Tzeng, Tsai, Chiang, Chang, & Nagai, 2013). Thus, S-P Chart can not only display the diagnostic assessment of learning, but also present the effectiveness of instructions (McArthur, 1983). That is to say, it has been adopted as a method for diagnosing, analyzing, processing and arranging data in a defined order (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013; You & Yu, 2006) so as to get the quality of test items and diagnosis of students' learning. The S-P chart gives matrix structure of student-problem described below (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013, p. 180):

$$\begin{aligned} P &= (p_1, p_2, \dots, p_n), \text{ then} \\ s_i &= (x_i(1), x_i(2), \dots, x_i(n)) \\ p_j &= (x_1(j), x_2(j), \dots, x_m(j)) \end{aligned} \quad (1)$$

where

$$x_i(j) = \begin{cases} 0, & \text{student } i \text{ gives wrong answer to problem } j \\ 1, & \text{student } i \text{ gives correct answer to problem } j \end{cases}$$

Then, caution index for students (CS) and caution index for items (CP) are as follows (Sheu, Chen, Tzeng, Tsai, Chiang, Chang, & Nagai, 2013, p. 18):

$$CS_i = 1 - \frac{\sum_{j=1}^n (y_{ij})(y_{kj}) - (y_{ik})(\bar{y})}{\sum_{j=1}^m (y_{kj}) - (y_{ik})(\bar{y})} \quad (2)$$

where  $\bar{y} = \frac{1}{n} \sum_{j=1}^n y_{kj}$  and  $m = y_{ik}$

$$CP_j = 1 - \frac{\sum_{i=1}^N (y_{ij})(y_{ik}) - (y_{kj})(\bar{y})}{\sum_{i=1}^m (y_{ik}) - (y_{kj})(\bar{y})} \quad (3)$$

where  $\bar{y}' = \frac{1}{N} \sum_{i=1}^N y_{ik}$  and  $m' = y_{kj}$

Based on the results of caution index for students (CS) and caution index for items (CP) above, the S-P chart can be drawn as shown in Table 1 (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013, p. 76):

Table 1. S-P chart

Student number	Problem number $P_j, j = 1, 2, \dots, n$	Total score	CS
$S_i, i = 1, 2, \dots, m$	$Y = (y_{ij})$	<div style="text-align: center;">           High  <math>\updownarrow SS_i</math>            Low         </div>	$CS_i$
Number of Correct answer	More $\longleftrightarrow$ Less $PP_j$	$\sum_{i=1}^m SS_i = \sum_{j=1}^n PP_j$	—
$CP$	$CP_j$	—	—

Sato (1980) and Sato and Kurata (1977) utilize student's caution index and their correct answer rate to divide the type of students' learning into six categories, and by comparing problem caution index and percentage of correct answers, question items are diagnosed and classified into six types as shown in Figure 1 (Sheu, Pham, Nguyen, & Nguyen, 2013, p. 76).

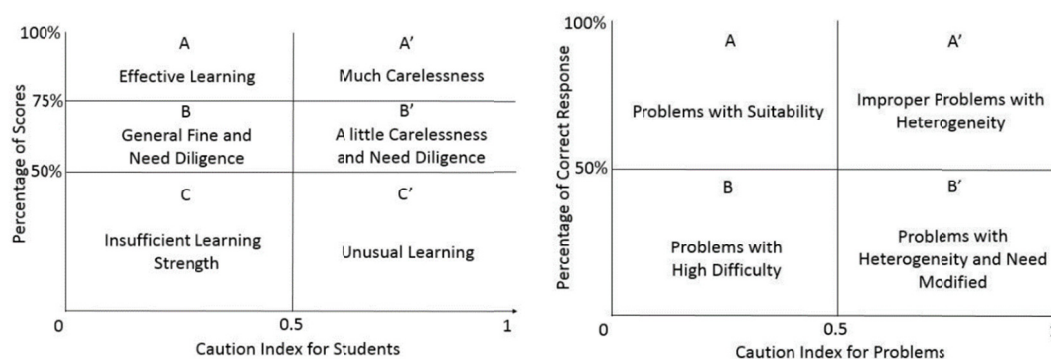


Figure 1. Type of students' learning (left) and question item (right)

### 1.2 Grey Relational Analysis (GRA)

Deng (1989) introduced the Grey system theory to deal with small and insufficient amount of information data, by using prediction and decision-making processes. It can transform the data from grey to white, and the satisfactory prediction results with partially unknown parameters can be obtained. Grey relational analysis (GRA) is one of the methods in Grey system theory which aims at processing uncertain, multiple, discrete and incomplete information effectively (Wen, Chou, Chang, Chen, & Wen, 2009). Thus, the GRA can not only count and quantize the discrete data, but also make them ordinal to be analyzed. The GRA procedures are described below (Sheu, Tzeng, Tsai & Chen, 2012).

1) Establishment of original vectors: The reference vector  $x_0$  and inspected vectors of original data  $x_i$  are calculated as follows (Sheu, Tzeng, Tsai, & Chen, 2012, p. 56):

$$\begin{aligned} x_0 &= (x_0(1), x_0(2), x_0(3), \dots, x_0(k), \dots, x_0(m)) \\ x_1 &= (x_1(1), x_1(2), \dots, x_1(k), \dots, x_1(m)) \\ x_2 &= (x_2(1), x_2(2), \dots, x_2(k), \dots, x_2(m)) \\ &\vdots \\ x_i &= (x_i(1), x_i(2), \dots, x_i(k), \dots, x_i(m)) \\ &\vdots \\ x_n &= (x_n(1), x_n(2), \dots, x_n(k), \dots, x_n(m)) \end{aligned} \quad (4)$$

2) Calculations of grey relational grade: Nagai, Yamaguchi and Li (2005) developed an equation to calculate the grey relation in which reference vector is  $x_0$  and comparative vector is  $x_i$ . When  $\gamma_{0i}$  is close to 1, it means that  $x_0$  and  $x_i$  are highly related to each other. On the other hand, if  $\gamma_{0i}$  is close to 0, the relationship between  $x_0$  and  $x_i$  is lower. The equation of the local grey relation is shown as follows (Sheu, Tzeng, Tsai, & Chen, 2012, p. 56).

$$\Gamma_{0i} = \Gamma(x_0(k), x_i(k)) = \frac{\bar{\Delta}_{\max} - \bar{\Delta}_{0i}}{\bar{\Delta}_{\max} - \bar{\Delta}_{\min}} \quad (5)$$

$$\text{where } \bar{\Delta}_{0i} = \|x_{0i}\|_{\rho} = \left( \sum_{k=1}^n [\Delta_{0i}(k)]^{\rho} \right)^{\frac{1}{\rho}}$$

When  $1 \leq \rho \leq \infty$ , it means Minkowski's grey relation; if  $\rho = 2$ , it means Euclidean grey relation. Then, the global grey relation is as follows (Sheu, Tzeng, Tsai, & Chen, 2012, p. 56).

$$\Gamma_{ij} = \Gamma(x_i, x_j) = 1 - \frac{\bar{\Delta}_{ij}}{\bar{\Delta}_{\max}} \quad (6)$$

$$\text{where } i, j \in 1, \dots, n, \quad \bar{\Delta}_{ij} = \left( \sum_{k=1}^n [\Delta_{ij}(k)]^2 \right)^{\frac{1}{2}}$$

3) Grey relation ordinal: The decision-making is based on the comparison of grey relation  $\Gamma_{0i}$ . It can be recognized as an important ordinal principle when  $\Gamma_{0i}$  number is larger. Thus, any vector can be the standard vector and others can become reference vectors. After calculating the grey relation, the grey relation matrix can be established to achieve the characteristic values so as to reach the grey relation ordinal. Sheu, Pham, Nguyen, and Nguyen (2013) suggest that through the ordinal, different causes can be identified, and the most influential cause can be found.

4) Calculation of grey relation: It is to normalize the data of reference vector for meeting three conditions: non-dimension, scaling and polarization. Three types of result will be found: larger-the-better, smaller-the-better and nominal-the-better (Yamaguchi, Li, & Nagai, 2005). Then, the outcome can be judged based on whether the expected goal has been reached.

### 1.3 Grey Student-Problem (GSP) Chart

By combining S-P chart and GRA, Nagai, Yamaguchi and Li (2005) establishes grey student-problem (GSP) Chart in order to overcome the insufficiency of the S-P chart. The purpose is to make the analysis more concrete

and accurate, so that uncertain factors can be dealt with and analyzed. The GSP chart matrix is shown in Table 2 (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013, p. 181).

Table 2. GSP Chart Matrix

Student number	Problem number $P_j, j = 1, 2, \dots, n$	Total score	LGRA-S
$S_i, i = 1, 2, \dots, m$	$X = [x_{ij}]_{m \times n}$	$SS_i$	High $\updownarrow$ $GS_i$ Low
Number of Correct answer	$PP_j$	$\sum_{i=1}^m SS_i = \sum_{j=1}^n PP_j$	—
$LGRA-P$	More $\longleftrightarrow$ Less $GP_j$	—	—

As can be seen,  $GS_i$  is the local grey relational grade of the  $i$ -th student, and  $GP_j$  is the local grey relational grade of the  $j$ -th problem. The calculations are shown below (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013, p. 181):

$$GS_i = \gamma_{0i} = \frac{\bar{\Delta}_{\max} - \bar{\Delta}_{0i}}{\bar{\Delta}_{\max} - \bar{\Delta}_{\min}}, i = 1, 2, \dots, m \quad (7)$$

$$GP_j = \gamma_{j0} = \frac{\bar{\Delta}_{\max} - \bar{\Delta}_{j0}}{\bar{\Delta}_{\max} - \bar{\Delta}_{\min}}, j = 1, 2, \dots, n \quad (8)$$

Based on  $GS_i$  ( $i = 1, 2, \dots, m$ ) value, each student is classified into five categories: A (Excellent,  $GS \geq 0.8$ ), B (Good,  $0.6 \leq GS < 0.8$ ), C (Medium,  $0.4 \leq GS < 0.6$ ), D (Weak,  $0.2 \leq GS < 0.4$ ) and E (Inferior,  $GS < 0.2$ ) (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013).

#### 1.4 Grey Structural Model (GSM)

Based on Grey theory and interpretive structural modeling (Nagai, Yamaguchi, & Li, 2005), Grey Structural Model (GSM) is developed to analyze the observed value of structural system, i.e., statistical graphical modeling. It calculated localized-GRA and globalized-GRA of grey relational grade and grey relational ordinal (Yamaguchi, Li, Akabane, & Kitaoka, 2007). Then, when a larger value appears, grey relational grade will be compared with each other, and the comparison will become the criterion of structural system arrangement. The formula of localized grey relational grade is (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013, p. 181):

$$\gamma_{oi} = \frac{\max_{\forall i} \|x_0 - x_i\|_{\rho} - \|x_0 - x_i\|_{\rho}}{\max_{\forall i} \|x_0 - x_i\|_{\rho} - \min_{\forall i} \|x_0 - x_i\|_{\rho}} \quad (9)$$

And formula of globalized grey relational grade is (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013, p. 181):

$$\gamma_{ij} = 1 - \frac{\|x_i - x_j\|_{\rho}}{\max_{\forall i} \max_{\forall j} \|x_i - x_j\|_{\rho}} \quad (10)$$

The hierarchical structure in GSM is to combine several relative elements as follows:

$$P = \{(x_i, x_j) | a_{ij} \geq \psi, a_{oi} < a_{oj}\} \quad (11)$$

Digraph can be drawn by using three parameters: distinguish coefficient  $\rho$  which decides the basic composition of digraph, class coefficient  $\theta$  which gives the hierarchy, and path coefficient  $\psi$  which gives an ordered pair of element arrows (Nguyen, Nguyen, Pham, Tsai, & Nagai, 2013).

### 1.5 Research Questions

This study aimed at evaluating the effect of a remedial instruction on fourth graders' English achievement and their learning condition. In order to examine students' learning improvement and test quality, this study adopted S-P chart to produce S-P diagrams for identifying the graphic position of each student and question item. Since GSP and GSM can be used to reveal students' ability clustering and the graphic structure of each cluster respectively, this study established GSP chart to understand students' performance and GSM structure to divide students into different clusters based on their performance. Hence, an English test was employed before and after one-semester remedial instruction and students' test scores were analyzed by establishing the S-P chart, GSP chart and GSM to answer the following research questions:

- 1) Is there a significant difference between English pre-test and post-test scores?
- 2) Are there differences in the S-P and GSP charts before and after remedial instruction?
- 3) To what extent do types of students' learning and question item change after remedial instruction?
- 4) To what extent do students' clusters in GSM graph change after remedial instruction?

## 2. Method

### 2.1 Participants

The participants of this study were twenty elementary fourth graders in Taipei City in northern Taiwan. All have been studying English as a school subject since first grade, and English is one of the subjects they all had to take from grade one to six. Based on their English test scores in the previous academic year, they were considered as underachievers who were required to take an hour remedial instruction every Wednesday afternoon after school at least for one semester.

### 2.2 Instrument

An English test, consisting of four sections with 25 question items in total, which was based on the textbook in the fourth grade, was developed by the participants' English teacher. The first section was phonics (6 items), followed by dialects in the section two (6 items). Section three focused on vocabulary (8 items), and the final section was about sentence structure, i.e., grammar (5 items). All the questions were multiple choices with only one correct answer, and 4 points were given to each correct answer which makes the total score 100.

The first draft of test was given to two colleagues who also taught English at the same school for examining the suitability of wording and question items so as to establish the expert validity of the test. Based on their comments, question items were revised to make them appropriate to the context and subjects in the study. Then, after conducting both tests, the scores were analyzed to establish their reliability and the Cronbach's alpha ( $\alpha$ ) value of the pre-test was .691, compared with .730 of the post-test. Disparity index (D) of both tests was .60. Accordingly, the two tests can be considered as having good internal consistency (George & Mallery, 2003).

### 2.3 Data Analysis

First of all, the numerical results of pre-test and post-test were written in a \*.xls file, and by SPSS/PC 18.0, paired samples *t*-test was first carried out to see whether there was significant difference between the two tests, i.e., significant level was set at .001. After that, the data were analyzed by using equations and Matlab 7.10 software. Students' scores were arranged from high to low, and the number of correct answer items was ordered from more to less so as to calculate all percentage, caution indices for students and items, and disparity index.

For sorting and calculating S-P chart, total S and P, S curve and P curve and disparity index were displayed in a S-P chart. Afterward student and item caution indices and correct answer rate were compared so that the type of students and items can be classified.

As to GSP chart, the equations of LGRA and Larger-the-better are used to calculate Gamma which can be used to generate a GSP chart. Students were arranged from high to low based on LGRA-S values, and item orders are from more to less depending on LGRA-P values. Then, the matrix of LGRA-S and LGRA-P values was calculated to produce GSP chart, total S and P, and GSP graphs.

As for GSM graph, the value of coefficients: distinguish coefficient  $\rho$ , class coefficient  $\theta$ , and path coefficient  $\psi$  were determined, followed by using Matlab toolbox for GSM to produce GSM structure diagrams so as to examine the relationship between students and test items.

### 3. Results

#### 3.1 Comparison of Pre-Test and Post-Test Scores

The test scores of pre-test and post-test were compared by pair-sample *t*-test and the result is shown in Table 3. As can be seen, there is a significant mean difference in the pre- and post-test ( $t = -17.899$ ,  $sig = .000$ ). Thus, it can be said that the students' English learning improve after receiving the remedial instruction.

Table 3. Paired samples *t*-test result of pre- and post-test

Test	Mean	SD	df	<i>t</i> -value
Pre-	66.0	15.654	19	-17.899***
Post-	84.8	12.755		

Note. \*\*\*  $p < .001$ .

#### 3.2 S-P Chart Analysis

The S curve and P curve of the pre- and post-test are shown in Figure 2.

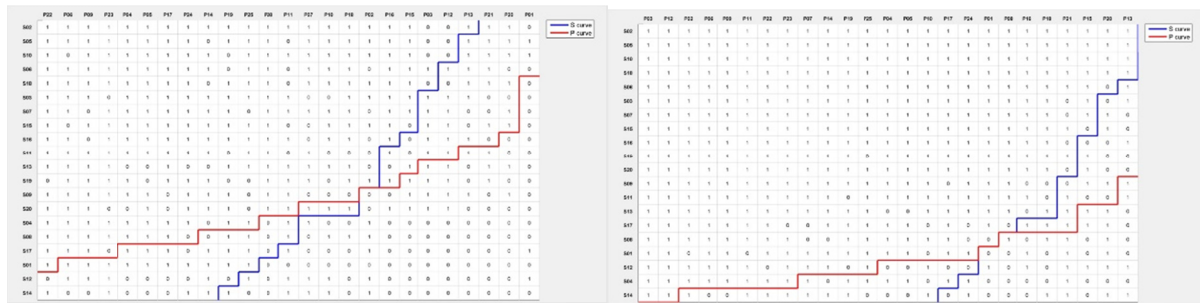


Figure 2. S-P chart for pre-test (and post-test)

If the caution index for student (CS) is greater than 0.5, the corresponding students and items should be paid attention to as shown in Table 4. In the pre-test, students  $S_2$  got the highest score while students  $S_{14}$  received the lowest one. When the level of question items is considered, item  $P_{22}$  was the easiest question, but item  $P_{01}$  was the most difficult one. In relation to CS, 12 students ( $S_5$ ,  $S_{10}$ ,  $S_7$ ,  $S_{15}$ ,  $S_{11}$ ,  $S_{13}$ ,  $S_{19}$ ,  $S_9$ ,  $S_{20}$ ,  $S_{17}$ ,  $S_{12}$ ,  $S_{14}$ ) needed teacher's attention, but it was reduced to 7 students ( $S_9$ ,  $S_{13}$ ,  $S_{17}$ ,  $S_8$ ,  $S_1$ ,  $S_{12}$ ,  $S_{14}$ ). Among them, student  $S_9$ ,  $S_{17}$ ,  $S_{12}$  and  $S_{14}$  should receive teacher's personal guidance in order to achieve better academic achievement.

Table 4. Student assessment results of the pre- and post-test

Pre-test						Post-test					
Student	Total	LGRG	Ratio	CS	Type	Student	Total	LGRG	Ratio	CS	Type
2	22	1	0.88	0.16	A	2	25	1	1	0	A
5	21	0.92	0.84	0.83*	A'	5	25	1	1	0	A
10	21	0.92	0.84	0.96*	A'	10	25	1	1	0	A
6	20	0.85	0.80	0.50	A'	18	25	1	1	0	A
18	20	0.85	0.80	0.50	A'	6	24	0.9	0.96	0.34	A
3	19	0.77	0.76	0.48	A	3	23	0.8	0.92	0.40	A
7	19	0.77	0.76	0.55*	A'	07	23	0.8	0.92	0.20	A
15	19	0.77	0.76	0.62*	A'	15	23	0.8	0.92	0	A
16	18	0.69	0.72	0.41	B	16	22	0.7	0.88	0.29	A
11	17	0.62	0.68	0.52*	B'	19	22	0.7	0.88	0.36	A
13	17	0.62	0.68	0.83*	B'	20	22	0.7	0.88	0.14	A
19	17	0.62	0.68	0.77*	B'	9	21	0.6	0.84	0.63*	A'
9	16	0.54	0.64	0.65*	B'	11	21	0.6	0.84	0.44	A
20	16	0.54	0.64	0.62*	B'	13	21	0.6	0.84	0.51*	A'
4	13	0.31	0.52	0.12	B	17	19	0.4	0.76	0.76*	A'
8	13	0.31	0.52	0.2	B	8	18	0.3	0.72	0.51*	B'
17	12	0.23	0.48	0.51*	C'	1	17	0.2	0.68	0.53*	B'
1	11	0.15	0.44	0	C	12	17	0.2	0.68	0.97*	B'
12	10	0.08	0.40	0.84*	C'	4	16	0.1	0.64	0.13	B
14	9	0	0.36	0.51*	C'	14	15	0	0.60	0.84*	B'

Note. \* CS > 0.5.

If the caution index for item (CP) is smaller than 0.5, the corresponding students and items should be paid attention to. Regarding CP (see Table 5), 14 items in the pre-test have been identified as having unclear meaning, with heterogeneous composition or not well prepared; but, the number in the post-test dropped to 6 items. Among them, item  $S_{23}$ ,  $S_7$ ,  $S_{19}$  and  $S_{25}$  should be revised to make them more suitable and effective.

Table 5. Item assessment results of the pre- and post-test

Pre-test						Post-test					
Student	Total	LGRG	Ratio	CS	Type	Student	Total	LGRG	Ratio	CS	Type
22	18	1	0.90	0.57*	A'	3	20	1	1	0	A
6	17	0.93	0.85	0.97*	A'	12	20	1	1	0	A
9	17	0.93	0.85	0.82*	A'	2	19	0.89	0.95	0.32	A
23	17	0.93	0.85	0.87*	A'	6	19	0.89	0.95	0	A
4	16	0.86	0.8	0.42	A	9	19	0.89	0.95	0	A
05	16	0.86	0.8	0.46	A	11	19	0.89	0.95	0.32	A
17	16	0.86	0.8	0.38	A	22	19	0.89	0.95	0.32	A
24	16	0.86	0.8	0.42	A	23	19	0.89	0.95	0.65*	A'
14	15	0.79	0.75	1.05*	A'	7	18	0.78	0.9	0.53*	A'
19	15	0.79	0.75	1.09*	A'	14	18	0.78	0.9	0.26	A
25	15	0.79	0.75	0.8*	A'	19	18	0.78	0.9	0.61*	A'
8	14	0.71	0.7	0	A	25	18	0.78	0.9	0.53*	A'
11	14	0.71	0.7	1.29*	A'	4	17	0.64	0.85	0.3	A
7	13	0.64	0.65	0.76	A'	5	17	0.64	0.85	0.32	A
10	13	0.64	0.65	0.67*	A'	10	17	0.64	0.85	0.26	A
18	13	0.64	0.65	0.13	A	17	17	0.64	0.85	0.32	A
2	12	0.57	0.60	1.06*	A'	24	17	0.64	0.85	0.38	A
16	12	0.57	0.60	0.03	A	1	16	0.56	0.8	0.05	A
15	11	0.5	0.55	0.16	A	08	15	0.44	0.75	0.04	A
3	10	0.43	0.5	0.58*	A'	16	15	0.44	0.75	0.35	A
12	10	0.43	0.5	0.58*	A'	18	15	0.44	0.75	0.39	A
13	9	0.36	0.45	0.16	B	21	15	0.44	0.75	1.22*	A'
21	9	0.36	0.45	0.16	B	15	13	0.22	0.65	0.35	A
20	8	0.29	0.4	0.21	B	20	13	0.22	0.65	1.06*	A'
1	4	0	0.2	1.11*	B'	13	11	0	0.55	0.62	A'

Note. \* CP > 0.5.



### 3.3 Types of Students' Learning

Caution index for students (CS) and percentage of scores were calculated to identify six types of students' learning (see Figure 3). The number of type A which represented effective learning, was increased threefold in the post-test, i.e., from 4 to 12 which represent 60% of the students; by contrast, a number drop was found in type A' (much carelessness) and B' (a little carelessness and need diligence) between pre-test and post-test. Three students ( $S_4$ ,  $S_8$ ,  $S_{16}$ ) were generally fine and needed diligence in learning (type B) in the pre-test, yet only one student ( $S_4$ ) remained in the same condition in the post-test. One student ( $S_1$ ) lacked of sufficient learning strength (type C) and three ( $S_{14}$ ,  $S_{12}$ ,  $S_{17}$ ) had unusual learning (type C') in the pre-test test, but they were not seen in the post-test. It should be noted that after the instruction, student  $S_4$  still stayed at the same type (type B) and student  $S_8$  declined from type B to type B'.

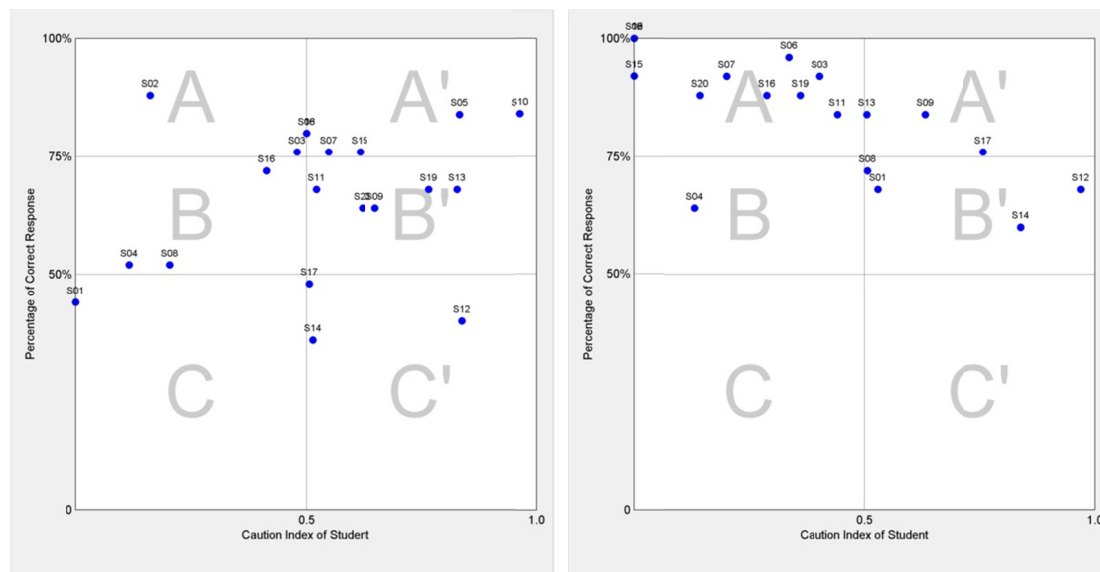


Figure 3. Student type diagram for pre-test (left) and post-test (right)

### 3.4 Types of Test Item

Caution index for problem (CP) and item passing ratio were analyzed to diagnose the question quality of pre-test and post-test, and question type diagrams are shown in Figure 4. As can be seen, there were 8 problems typed A in the pre-test, which were suitable for distinguishing low achievers from other types of students; however, the number showed a tremendous increase of 18 questions in the post-test. By contrast, the number in the other three types dropped; that is, for type A', 7 questions contained improper problems with heterogeneity, but the number was decreased slightly to 5 ( $P_7$ ,  $P_{13}$ ,  $P_{19}$ ,  $P_{23}$ ,  $P_{25}$ ). Also, 3 questions ( $P_{20}$ ,  $P_{21}$ ,  $P_{13}$ ) were very difficult (type B) and 2 questions ( $P_3$ ,  $P_{12}$ ) were high heterogeneous and need modification (type B') which has high difficulty, but no question appeared in both types in the post-test. It should be pointed out that item  $P_7$ ,  $P_{23}$  and  $P_{25}$  remained at the same difficult level (A') as in the pre-test, and item  $P_{20}$  and  $P_{21}$  which were typed B in the pre-test, cannot be categorized into any of the four types in the post-test.

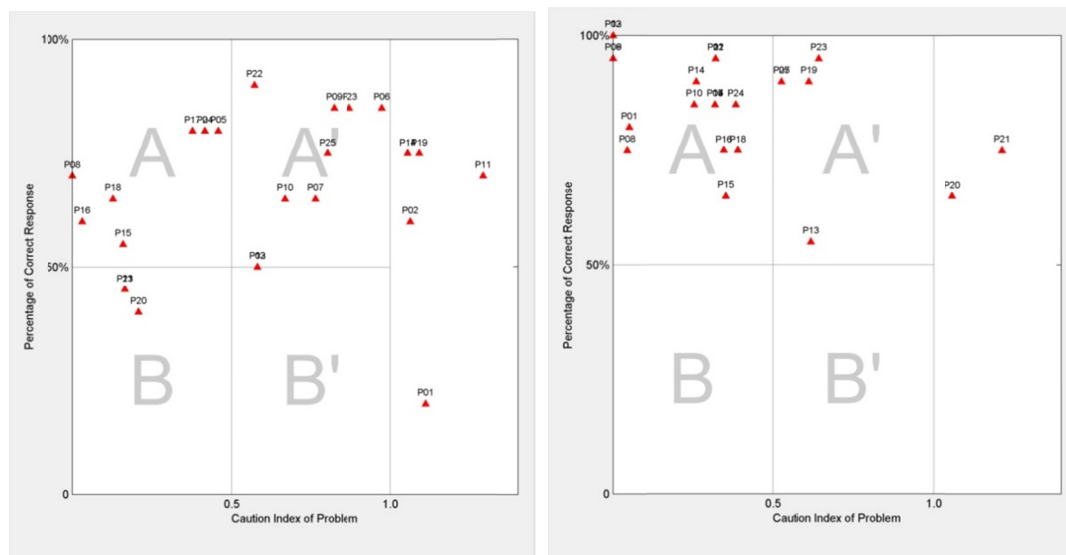


Figure 4. Question type diagram for pre-test (left) and post-test (right)

### 3.5 GSP Graph Analysis

Students' scores of pre- and post-test were analyzed, and the results were classified into five categories as shown in Table 6.

Table 6. The classification results of students

GS value	Category (Achievement)	Pre-test	Post-test
$GS \geq 0.8$	A (Excellent)	5	8
$0.6 \leq GS < 0.8$	B (Good)	7	6
$0.4 \leq GS < 0.6$	C (Medium)	2	1
$0.2 \leq GS < 0.4$	D (Weak)	3	3
$GS < 0.2$	E (Inferior)	3	2

The GSP graph of pre- and post-test was established by using a combination of the S-P chart analysis and the results of GSP chart analysis as shown in Figure 5. It should be pointed out that the intersection point of Logistic Regression equations and Rasch line = 0.5 represent the mean value and item difficulty of the two tests. As can be seen, the students' mean value fell slightly from 38.5% to 37.5% which means students' gamma value did not increase obviously. The Figures also show that the item difficulty was reduced by 9% (from 30% to 21%). This means that the question items in the post-test became easier than those in the pre-test. Because of this, it can be said that that the remedial instruction had a positive effect on students' learning.

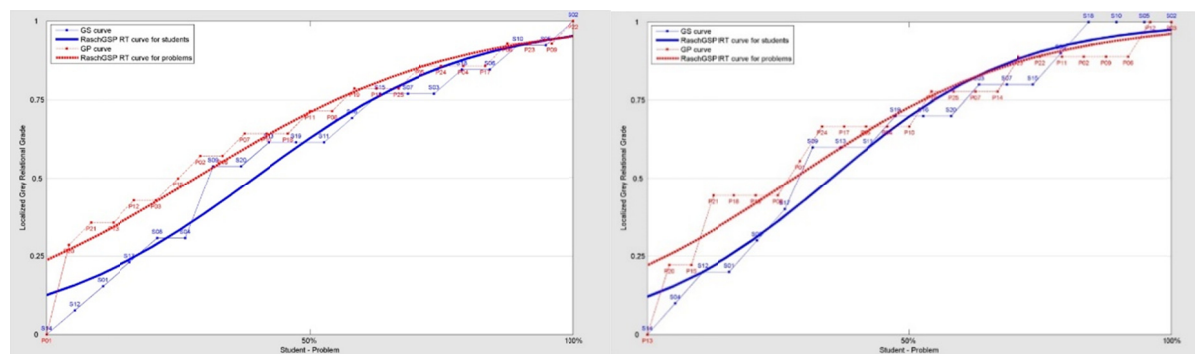


Figure 5. GSP graph of pre-test (left) and post-test (right)

### 3.6 Cluster Analysis of GSM Structure

GSM structural analysis is run through a matrix using Matlab software to produce the GSM structural graph of the students and items as describes below.

#### 3.6.1 GSM Structural Graph of Students

The GSM structural graph of students of the pre- and post-test are shown in Figure 6. Both graphs have 6 layers, that is, while the students at the top of the diagram were high achievers with higher English test scores, those at the bottom are underachievers with lower English test scores. The hierarchical structure in GSM combined several relative elements and grey relational paths existed among students, as follows.

##### 1) Grey relational grade ( $\gamma$ ) > 0.75

- Pre-test:  $S_2 \succ S_5, S_{10} \succ S_6, S_{18}$
- Post-test:  $S_2, S_5, S_{10}, S_{18}$

##### 2) Grey relational grade ( $\gamma$ ) $\leq 0.50$

- Pre-test:  $S_{14}, S_{12}, S_1 \prec S_{17}, S_8, S_4 \prec S_{20}, S_9, S_{19}$
- Post-test:  $S_{14}, S_4, S_{12}, S_1 \prec S_{17}, S_8 \prec S_{13}, S_{11}, S_9 \prec S_{20}, S_{19}, S_{16}$

After the remedial instruction, 4 students ( $S_2, S_5, S_{10}$  and  $S_{18}$ ) with grey relational grade ( $\gamma$ ) > 0.75 reached at the top layer, compared with only one student ( $S_2$ ) before receiving remedial instruction. However, when grey relational grade ( $\gamma$ )  $\leq 0.50$  is considered, the number of students in the post-test showed an increase of 4 and scattered over 3 layers.

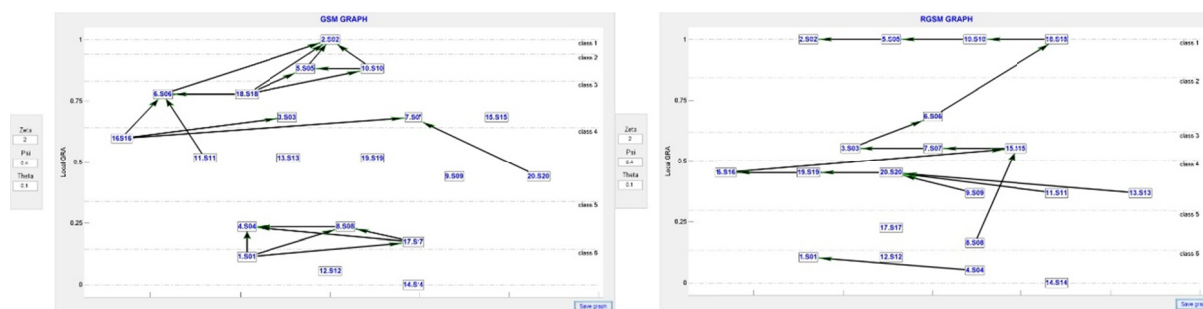


Figure 6. Students' GSM structural graph of pre-test (left) and post-test (right)

#### 3.6.2 GSM Structural Graph of Question Items

The GSM structural graph of question items of the pre- and post-test are shown in Figure 7. Both graphs have 6 layers, and items on the top layer represent the most difficult, but the lowest layer has just the reverse. The hierarchical structure combined several relative elements can be seen and the following lists grey relational paths in both graphs.

1) In order to evaluate the effects of remedial instruction effectively, the grey relational paths of the post-test were investigated in particular. A close relationship was found among  $P_8, P_1$  and  $P_{24}$ , and as having the highest correct rate,  $P_1$  and  $P_{24}$  can be deleted. For the same reason,  $P_{10}, P_{23}$ , and  $P_{25}$  can also be removed from the test.

2) When grey relational grade ( $\gamma$ ) > 0.75 is taken into account, 8 items ( $P_{22}, P_6, P_9, P_{23}, P_4, P_5, P_{17}, P_{24}$ ) were in accordance with this grade in the pre-test, but only 2 items ( $P_3, P_{12}$ ) reminded the same grade in the post-test.

3) Taken account of grey relational grade ( $\gamma$ )  $\leq 0.50$  and > 0.25, 7 items ( $P_{21}, P_{13}, P_{12}, P_3, P_{15}, P_{16}, P_2$ ) were found in the pre-test, but there was 10 items ( $P_{21}, P_{18}, P_8, P_{16}, P_1, P_{24}, P_{17}, P_{10}, P_5, P_4$ ) in the post-test. Taken remedial instruction into consideration, these question items play key roles in promoting the effects of students' achievement.

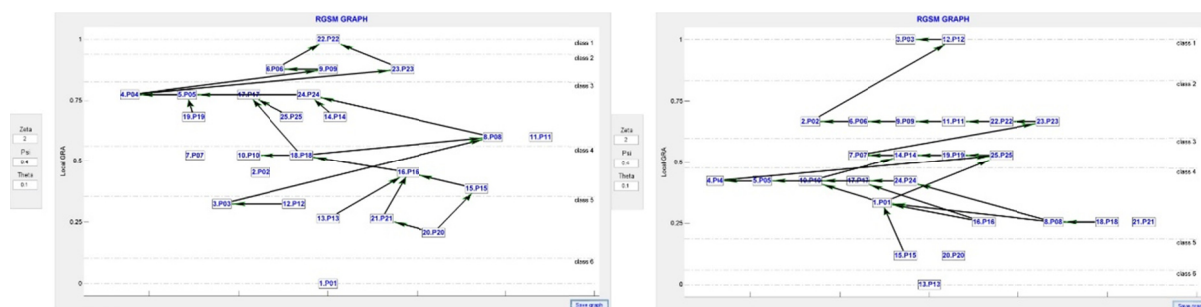


Figure 7. Question items' GSM structural graph of pre-test (left) and post-test (right)

#### 4. Discussion

Most statistic methods (i.e., *t*-test, correlation and ANOVA) are commonly used to identify the impact of teaching programs on students' learning, as in this research in which after implementing a remedial instruction, a considerable improvement of students' achievement was found by paired samples *t*-test. This is actually in keeping with the purposes of statistics analysis in which the results and implications can be broader and generalized. In fact, it is necessary to note that a key criticism of existing statistics is that they often lead to more general and overall results rather than more specific ones. More importantly, when remedial instruction in which the progress of each student is emphasized is concerned, this is one of the limitations in most statistic methods, which in part, is for the lack of precise design on the part of statistic techniques. This is one of the key motivations for initiating this research that the statistics methods based on Grey theories can effectively assess individual students' achievement after a remedial instruction.

In this paper, the S-P chart, GSP chart and GSM method were used to examine the effect of a remedial instruction on fourth grade students in Northern Taiwan. The aim was to demonstrate these methods can be used as useful and effective tools within the classroom for elementary English teachers to measure individual students' learning condition and achievement. The findings and implications of this research are as follows.

First of all, from S-P chart, S-curve and P-curve visually represented the average percentage of correctly answered tests, which mean showing students' correct answers rate and test's item difficulty. Both can be used to determine whether the questions have the discrimination, and this can be applied to estimate the students' learning which needs to be strengthened. It is true to say that the degree of appropriateness and effectiveness of teaching is limited if students' test scores are low, and students' learning achievement will not become evident if the item difficulty had been high.

Another way of analyzing caution index for students (CS) and for items (CP) is the classification of students' learning condition and item difficulty. Both classifications contained 6 types depict the students' movement from ineffective to effective and the item changes from difficult to easy, respectively. As with the nature of Type B (much carelessness) and B' (a little carelessness and need diligence), the attitudinal element in learning is probably the most critical, and as such, attitudes have been suggested as the starting point in many learning situations. There are, of course, instructional adaptations (e.g., content, materials, activities, etc.) based on students' conditions that can be made, and this will remain to future research to probe the effect of adaptations on students' learning condition and performance.

In addition, the GSP chart presented in the paper can evaluate each student's improvement in an English remedial instruction by comparing the difference between pre-test and post-test. The difference of GSP graphs between pre-test and post-test indicated that the effect of remedial instruction is shown in the changes of the students' average score and a fall of the item difficulties from Rasch line = 0.5. The slope of the lines in graphs not only clarifies uncertain aspects of their learning but also suggests areas for each student to be improved.

Finally, the GSM method draws the diagram which shows a relationship among students and question items with a directive path and their position with a visual hierarchy. That is to say, it can be used to establish a learning map of each individual student, so that teachers can know their students' weaknesses to be deal with in class so as to promote their teaching efficiently. And for students, they can be taught based on their language ability, and this can increase their confidence and then the progress of their learning is more likely to be seen.

A limitation of this research is that there was an elementary school English remedial instruction to verify the clustering and hierarchy, but under such a small amount of number and item number, the study provides teachers

with an evaluation means of their teaching and students' learning during and after teaching. Given that there is no official assessment instrument to be used in reference to remedial programs in elementary schools in Taiwan, it will be important for instructors to at least be aware of different evaluation methods instead of merely using school tests, and then adopt them necessarily and effectively.

Another implication is that to assist in assessing remedial instruction in this study, a combination of different graphic analysis seems to be relevant to elementary school English teachers. That is, based on the graphic results, it can be used at the middle of their teaching in order to make any adjustment for the rest of instruction, or at the end as a review for the effectiveness of instruction or implications for the succeeding lessons to keep their teaching efficient and successful. It is obvious that this study's findings, along with the result of previous research, benefit elementary school English teachers and instructors of remedial programs especially in assessing their students' progress or achievement of English language learning.

To sum up, the combination of S-P chart, GSP chart and GSM method in this study served as an assessment system to produce graphic results from a small amount of students and test items. It appears to be an effective method of clustering students' learning achievement and condition and a practical way to find out students' learning progress in graphic ways. This can be used to serve as a precise learning map for teachers and for instructors in remedial programs in particular, that is to say, it provides useful information about designing and organizing lessons and then assigning different tasks, activities and assignments to different groups based on the clustering result. Consequently, teachers can adjust their teaching to be more appropriate and beneficial to their students.

## 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, NY: Center for the Analysis of Postsecondary Readiness.
- Chen, D. J., Lai, A. F., & Jhan, Y. W. (2007). *The correlation analysis of the test item attributes and the test-takers' performance in web-based test* (pp. 435–442). Proceeding of the IADIS International Conference e-learning.
- Chen, D. J., Lai, A. F., & Liu, I. C. (2005). The design and implementation of a diagnostic test system based on the enhanced S-P model. *Journal of Information Science and Engineering*, 21, 1007–1030.
- Csizer, K., & Magid, M. (2015). *The impact of self-concept on language learning*. Bristol, UK: Multilingual Matters. <https://doi.org/10.21832/9781783092383>
- Deng, J. L. (1989). Introduction to grey system theory. *The Journal of Grey System*, 1(24), 1–24.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Boston, MA: Allyn & Bacon.
- Greig, A., Taylor, D. J., & MacKay, T. (2012). *Doing research with children: A practical guide*. London, UK: Sage. <https://doi.org/10.4135/9781526402219>
- Jacobs, G., & Farrell, T. S. C. (2001). Paradigm shift: Understanding and implementing change in second language education. *TESL EJ*, 5(1). <http://www.kyoto-su.ac.jp/information/teslej/ej17/toc.html>
- Jones, J., & Coffey, S. (2016). *Modern foreign languages 5–11: A guide for teachers*. Oxford, UK: Taylor and Francis.
- Kirsch, C. (2008). *Teaching foreign languages in the primary school*. London, UK: Continuum.
- Lin, Y. H., & Chen, S. M. (2006). The integrated analysis of S-P chart and ordering theory on equality axiom concepts test for sixth graders. *Wseas Transactions on Mathematics*, 5(12), 1303–1308.
- McArthur, D. L. (1984). *Analysis of test score patterns: The student-problem (S-P) technique*. Oakland, CA: University of California. <https://doi.org/10.1037/e679012011-001>
- Ministry of Education. (2003). *Grade 1–9 curriculum guidelines*. Taipei, TW: Ministry of Education.
- Mitchell, R., & Myles, F. (2004). *Second language learning theories*. New York, NY: Hodder Arnold.
- Nagai, M., Yamaguchi, D. S., & 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. *Journal of Taiwan*

- Kansei Information*, 4(4), 179–190.
- Pham, D. H., Sheu, T. W., Tsai, C. P., Nguyen, P. T., Nguyen P. H., & Nagai, M. (2013). *New application based on GRA and Rasch model GSP in evaluating academic achievement*. 2013 Kansei Information of Creative Contest, Chien Kuo Technology University, Taichung, Taiwan.
- Salimi, E. A., & Farsi, M. (2018). An investigation of assessment literacy among native and nonnative English teachers. *Journal of English Language Teaching and Learning*, 10(22), 49–62.
- Sato, T. (1969). *A method of analyzing data gathered by the response analyzer for diagnosis of student performance and the quality of instructional sequence*. Proceedings of IECE of Japan annual conference, S12-1.
- Sato, T. (1980). The S-P chart and the caution index. *NEC Educational Information Bulletin*, 80(1), 445–450. [https://doi.org/10.1016/0045-8732\(80\)90036-4](https://doi.org/10.1016/0045-8732(80)90036-4)
- Sato, T., & Kurata, M. (1977). Basic S-P score table characteristics. *NEC Research and Development*, 47, 64–71.
- Sheu, P. H. (2015). Examining university students' motivation and their motivational behaviors in English learning with Structural Equation Modeling. *International Journal of Humanities and Social Science*, 5(8), 21–29.
- Sheu, P. H. (2016). An analytic hierarchy process for English learning strategies of college students in Taiwan. *International Journal of Language and Linguistics*, 3(5), 30–39.
- Sheu, P. H. (2017). Examining the relationship of motivation, attitude, anxiety and achievement in English learning among elementary school students in Taiwan. *International Journal of Language and Literature*, 5(2), 174–184. <https://doi.org/10.15640/ijll.v5n2a18>
- Sheu, P. H. (2019). Analyzing relationships among activities for teacher development, teaching beliefs, efficacy and teaching behaviors of elementary school English teaching in Taiwan. *International Research in Higher Education*, 4(3), 1–16. <https://doi.org/10.5430/irhe.v4n3p1>
- Sheu, T. W., Chen, T. L., Tzeng, J. W., Tsai, C. P., Chiang, H. J., Chang, C. L., & Nagai, M. (2013). Applying misconception domain and structural analysis to explore the effects of the remedial teaching. *Journal of Grey System*, 16(1), 17–34.
- 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), 75–86.
- Sheu, T. W., Tzeng, J. W., Liang, J. C., Wang, B. T., & Nagai, M. (2012). A study on Rasch model GSP chart and grey structural modeling. *Journal of National Taichung University: Mathematics, Science & Technology*, 26(1), 1–25.
- Sheu, T. W., Tzeng, J. W., Tsai, C. P., & Chen, T. L. (2012). Applying problem-concept chart combined with structural analysis to investigate the learning misconcept-simple equation with one variable for example. *Journal of Grey System*, 15(1), 55–66.
- Tsai, P. Y., & Chung, C. J. (2003). The study of applying interpretive structural modeling in instructional structural design. *Educational Research & Information*, 11(2), 1–40.
- Wang, B. T., Sheu, T. W., Liang, J. C., Tzeng, J. W., & Nagai, M. (2011). The integrated methods of GSP and GSM in concepts diagnosis for English grammar. *International Journal of Kansei Information*, 2(2), 87–100.
- Wen, K. L., Chou, C. S., Chang, H. C., Chen, X. Y., & Wen, H. C. (2009). *Grey system and applications*. Taipei, TW: Wunan.
- Yamaguchi, D. S., Li, G. D., Akabane, K. M. T., & Kitaoka, M. (2007). A realization algorithm of grey structural modeling. *Journal of Grey System*, 10(1), 33–40.
- Yamaguchi, D. S., Li, G. D., & Nagai, M. (2005). New grey relational analysis for finding the invariable structure and its applications. *Journal of Grey System*, 8(2), 167–178.
- Yamaguchi, D. S., Li, G. D., & Nagai, M. (2007). Verification of effectiveness for grey relational analysis models. *Journal of Grey System*, 10(3), 169–181.
- You, S. C., & Yu, M. N. (2006). The relationships among indices of diagnostic assessments knowledge structures and S-P chart analysis. *Education and Psychology*, 29(1), 183–208.

**Copyrights**

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).