

The Design and Implementation of a Diagnostic Test System Based on the Enhanced S-P Model

DENG-JYI CHEN, AH-FUR LAI AND I-CHANG LIU
Department of Computer Science and Information Engineering
National Chiao Tung University
Hsinchu, 300 Taiwan
E-mail: djchen@csie.nctu.edu.tw

The cognitive diagnostic test can be used to understand the learning effects (such as strengths and weaknesses) of learners for a specific subject area. Based on the evaluation results of the diagnostic test, instructors may suggest or give students additional material on the subject area for students who do not meet the requirements. The S-P (Student-Problem) model has been used for this purpose for a long time. However, the current S-P model pays little attention to a student's response time for each test item or for the entire question set during the test. The student's response time for each test item can be an important factor for instructors wanting to diagnose each student's individual ability in problem solving. Also, there are few computerized diagnostic test analysis systems available that are designed to support both text-diagram-based and/or multimedia-based presentation test items.

In this research, we incorporate the response time, difficulty index, and discriminatory index of each test item into an S-P model during the analysis. Specifically, we employ two terms: 1) the nimbleness of thinking of a student, which can be measured based on the response time for answering each test item, and 2) the problem solving ability of a student, which can be measured based on the student's ability to solve adaptive type questions with various difficulty levels and discrimination powers. With the incorporation of these parameters, an enhanced S-P model is presented. It can be used to diagnose both time dependent information and problem solving abilities with respect to test items and test-takers. A web-based computerized system was designed and implemented based on the enhanced S-P model for both text-diagram type presentation test items and multimedia type presentation test items. Practical examples were investigated and experimental studies conducted using the cognitive diagnostic computerized system to demonstrate the rationality and applicability of the proposed enhanced S-P model.

Keywords: enhanced S-P model, diagnostic test system, nimbleness of thinking, test-item solving ability, response time

1. INTRODUCTION AND LEGEND

1.1 Introduction

In the context of education, one possible purpose of a test is to assess whether an individual learner has attained an educational goal. The outcome of such a test can help instructors analyze problems with his or her instruction and to better understand the learner's strengths and weaknesses in a given subject. Furthermore, it can be used to fine-tune the instruction environment or method, based on the analytical results, to im-

Received July 4, 2004; revised March 15, 2005; accepted May 10, 2005.
Communicated by Robert Lewis.

prove the instructor's teaching performance. There are currently two methods for conducting tests. The traditional method of using paper and pencil tests, including the creation of test items, the grading of students' test sheets, and the analysis of test-takers' responses for each test item, is considered to be tedious [1]. The other method, Computer-Based Tests (CBT), allows test activities to be carried out in a computer environment. CBTs have several advantages over traditional methods; for example, they provide multimedia test questions and shorten test times. With the widespread use of the personal computers and the Internet, Web-Based Instruction (WBI) has also gained popularity and been adopted in distance learning and teaching. In WBI, assessing and diagnosing a learner's strengths and weaknesses is an important issue. Currently, the CBT is considered the main approach to evaluating the performance and capabilities of online learners in a WBI environment.

Diagnosis is an important part of the learning process. It helps instructors 1) analyze the abnormal performance of learners, 2) understand the suitability of given test items for test-takers, 3) evaluate test-takers' responses to each test item, and 4) pinpoint misconceptions of learners with respect to a given test subject. All of these issues must be investigated so that instructors can improve their instruction methods. Most of the current CBT systems support only fundamental diagnosis information, such as difficulty indexes and discrimination indexes, in regard to test items. The S-P chart (Student-Problem chart), which was originally proposed by Sato [6, 7, 10, 12-14], has been used to solve some of the above-mentioned problems. It provides caution indexes for diagnosing the abnormal performance of examinees, and a disparity coefficient for diagnosing test item sheets. However, the current S-P chart neglects time-dependent issues associated with test items and test-takers. In this paper, we propose two indexes, 1) the nimbleness of thinking, which is time dependent, and 2) the problem solving abilities of test-takers, in order to enhance the S-P model. The proposed model was integrated into an existing web-based CBT system. Practical examples were investigated and experimental studies conducted using the proposed cognitive diagnostic computerized tool to demonstrate the rationality and applicability of the proposed enhanced S-P model.

1.2 Legend

To help readers to follow our presentation, the following notations are defined.

S-P score table: the student-problem score table;

D^* : the disparity index, which is the ratio of the diverse area of the practical response to the expectation area of the random condition between the S curve and P curve;

C : the total numbers of "1's" and "0's" in the area surrounded by the S curve and P curve;

N : the total number of students taking the test;

n : the number of test items;

\bar{p} : the average ratio of correct responses;

$DB(M)$: the constant array that can be accessed by M ;

M : $M = \text{Gussan}(\sqrt{Nn} + 0.5)$;

CS_i : the student caution index;

CP_i : the item caution index;

- y : a two-dimensional array that stores test-takers' response patterns, which are sorted in descending order according to both the percentage of correct responses to test items and the percentage of correct responses of students;
- y_{ij} : the i^{th} person's response for item j ;
- y_i : the number of correct responses of student i on all of the test items;
- y_j : the number of correct responses for all of test-takers for test item j ;
- μ : the average score of the students;
- μ' : the average number of correct responses for all test items;
- $Tmax(j)$: the maximum response time of all test-takers for test-item j ;
- $Tmin(j)$: the minimum response time of all test-takers for test-item j ;
- $Wun(j)$: the weighted unit value, which is the distance between $Tmax(j)$ and $Tmin(j)$;
- $W(i, j)$: the weighted nimbleness value of test-taker i for item j ;
- $Tavg(j)$: the average response time of all test-takers for item j ;
- $Tlen(i, j)$: the response time of test-taker j ;
- $Nim(i)$: the nimbleness of thinking of test-taker i for all test items (from 1 to N), which can then be defined as the summation of the weighted nimbleness values of all items;
- $p(j)$: the difficulty index of test item j ;
- $Wp(j)$: the weighted value of the difficulty index of test item j ;
- $D(j)$: the discriminative index of test item j ;
- $Wd(j)$: the weighted value of the discriminative index of test item j ;
- $Ws(j)$: the weighted value of the problem solving ability for test item j ;
- $Right(i, j)$: the vector that records the correct response of item j for test-taker i ;
- $Sol(i)$: the test-item solving ability of test-taker i .

2. RELATED WORKS

The S-P chart was originally proposed by Sato in 1975 [12-14], and it has since been used to diagnose the aberrant performance of students' and to analyze the suitability of test items by means of an item caution index, a student caution index, and a disparity coefficient. The S-P chart does not make any assumptions about the test subject group and is based on a nonparametric statistical approach. It is suitable for application to formative tests used in classrooms. Through the analysis of students' response patterns, the caution indexes and disparity indexes can be calculated. To help teachers interpret these indexes, diagnostic information can be obtained and used as guidance for providing additional instruction for learners after an examination.

The S-curve shows how students agree with the problems, and the P-curve shows how the problems agree with the students. In an ideal situation, these two curves should coincide, but in a practical situation, these two curves will diverge [9, 11]. The disparity index (D^*) is the ratio of the divergence area of the practical response to the expectation area of the random condition between the S curve and P curve. Sato proposed an approximated formulation which is restated below:

$$D^* = \frac{C}{4Nn\bar{p}(1-\bar{p})D_B(M)}. \quad (1)$$

Here, C is the dimension between the S curve and P curve, i.e., the total amount of “1’s” and “0’s” in the area surrounded by the S curve and P curve. N is the total number of students, n is the number of test items, \bar{p} is the average ratio of correct responses, and $DB(M)$ is the constant array, which be accessed by M (where $M = \text{Gussan}(\sqrt{Nn} + 0.5)$).

The value range of the caution index is between 0 and 1. The higher the student’s caution index value, the greater is the abnormality of the student’s response. The higher the item caution index, the worse is the suitability of the item. The item caution index and the student caution index are calculated using Eqs. (2) and (3), respectively:

$$CS_i = 1 - \frac{\sum_{j=1}^n (y_{ij})(y_{.j}) - (y_i)(\mu')}{\sum_{j=1}^{y_i} y_{.j} - (y_i)(\mu')} = \frac{\sum_{j=1}^{y_i} (1 - y_{ij})(y_{.j}) - \sum_{j=y_i+1}^n (y_{ij})(y_{.j})}{\sum_{j=1}^{y_i} y_{.j} - (y_i)(\mu')}, \quad (2)$$

$$CP_j = 1 - \frac{\sum_{i=1}^N (y_{ij})(y_i) - (y_{.j})(\mu)}{\sum_{i=1}^{y_j} y_i - (y_{.j})(\mu)} = \frac{\sum_{i=1}^{y_j} (1 - y_{ij})(y_i) - \sum_{i=y_j+1}^N (y_{ij})(y_i)}{\sum_{i=1}^{y_j} y_i - (y_{.j})(\mu)}, \quad (3)$$

where y is a two-dimensional array that stores test-takers’ response patterns, which are sorted in descending order according to both the percentage of correct responses to test items and the percentage of correct responses of students (as illustrated in Fig. 1); y_{ij} represents the i^{th} person’s response for item j ; y_i represents the number of correct responses of student i for all test items; y_j represents the number of correct responses of all test-takers for test item j ; μ represents the average score of the students; and μ' represents the average number of correct responses for all test items.

The diagnostic result for each test item is divided into four parts according to the item caution index and the students’ correct response percentage for that test item [10-14]. The four parts (A, A', B, B') can be used to explain the possible outcomes of test items from the diagnosis as shown in Fig. 2. Part A shows that the test item is suitable and can be used to discriminate both the low performance student group and different types of student groups. Part B indicates that the item can be used to discriminate the high performance student group and other types of student groups. Part A' shows that the test item perhaps contains some abnormal components or that the test item includes some inadequate choices. Thus, part of the content of the test item must be modified. Part B' indicates that the test item is poor or that the semantics of the test item are ambiguous.

The diagnostic results for the student’s learning performance is divided into six parts, based on the student’s caution index and the correct response percentage of the student. These parts are shown in Fig. 3. Part A indicates that the student has good learning performance and high stability in testing. Part B shows the middle stability of the student, indicating that the student should work a bit harder. Part C indicates that the student’s learning capability is poor and that his or her mastery of the content is too low,

or that there is a need for work hard. Part A' shows that the student has good performance on tests but sometimes gives incorrect responses due to carelessness. Part B' indicates that the student's mastery of the content is not good, signifying that the student sometimes gives incorrect responses due to carelessness, and that he or she needs to work harder. Part C' shows that the student has low stability, and that his or her mastery of the content is very poor.

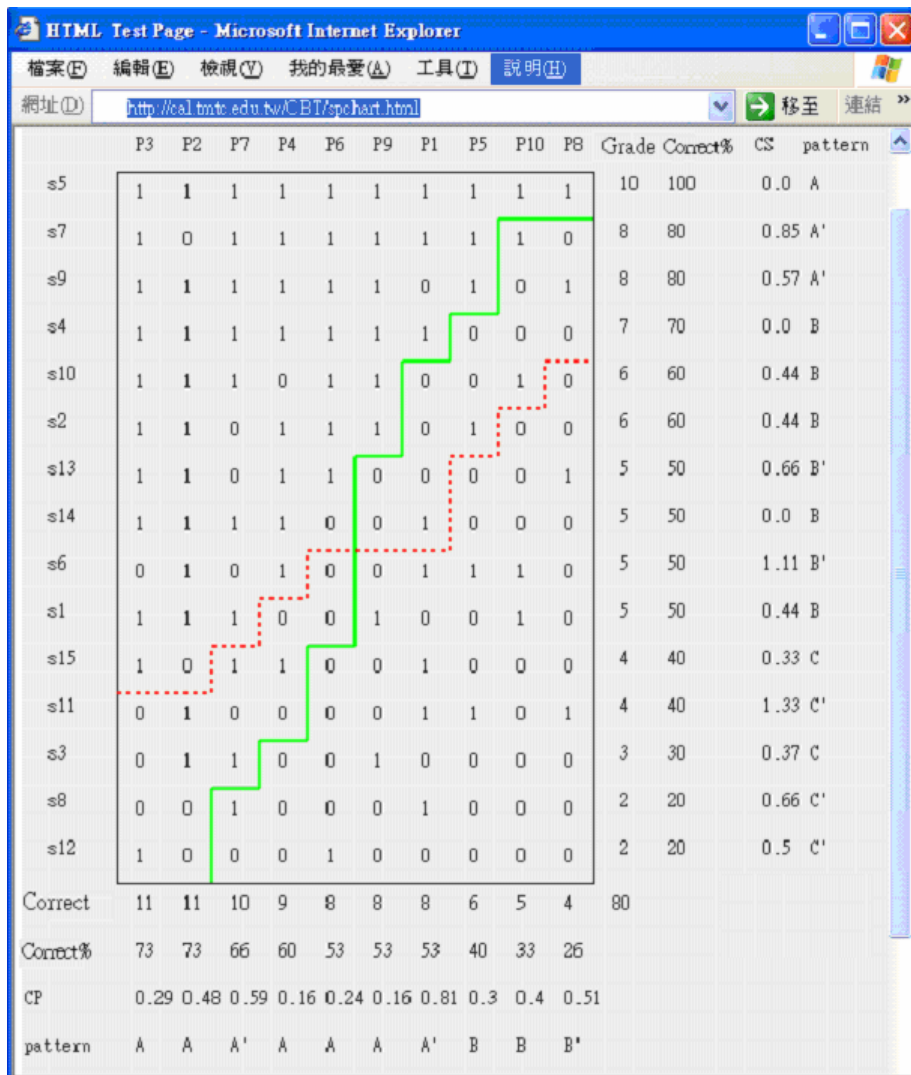


Fig. 1. Example of an S-P chart (straight line: S curve; dotted-line: P curve).

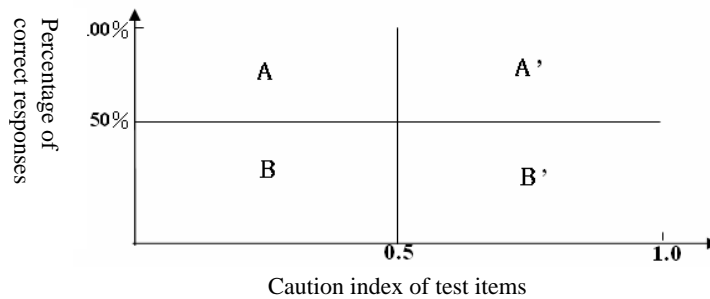


Fig. 2. Item categories based on the item caution index and item correct response rate.

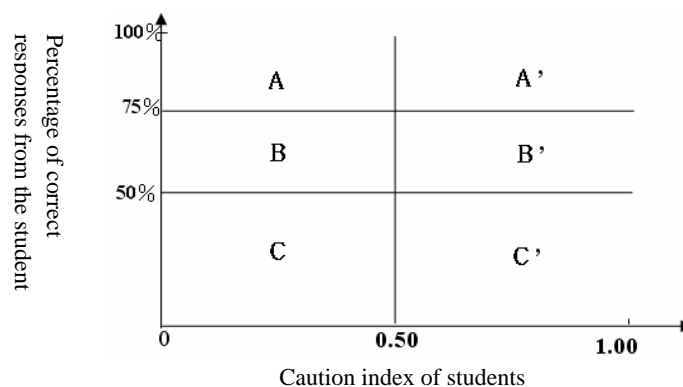


Fig. 3. Student's learning performance categories based on the student caution index and the student's correct response rate.

3. PROBLEM STATEMENTS AND PROPOSED SOLUTIONS

Previously, the S-P score table proposed by Sato did not take time factors into consideration. This work incorporates the student's response time, difficulty index, and discriminatory index for each test item. Specifically, we add and employ two terms: 1) the nimbleness of thinking of the student, which can be measured based on the student's response time in answering each test question; and 2) the student's ability to solve each test item, measured based on the student's ability to answer higher difficulty level questions correctly, and on a discriminatory index. With the incorporation of these parameters, an enhanced S-P model is presented. It can be used to diagnose both time dependent information and problem solving abilities with respect to test items and test-takers.

3.1 The Nimbleness of Thinking

The response time of the test-taker during online testing is an important factor that can be used to understand the speed of thinking of the test-taker for each test item (which is called the nimbleness of thinking, *Nim*). It is formulated as explained below.

$Tmax(j)$ is defined as the maximum response time of all test-takers for test-item j . $Tmin(j)$ is defined as the minimum response time of all test-takers for test-item j . $Wun(j)$ is defined as the weighted unit value, which is the distance between $Tmax(j)$ and $Tmin(j)$:

$$Wun(j) = (Tmax(j) - Tmin(j))/5. \tag{4}$$

The weighted nimbleness value of test-taker i for item j , $W(i, j)$, is formulated as in (5), where $Tavg(j)$ is the average response time of all test-takers for item j and $Tlen(i, j)$ is the response time of test-taker j :

$$W(i, j) = (Tavg(j) - Tlen(i, j))/Wun(j). \tag{5}$$

The nimbleness of thinking of test-taker i for all test items (from 1 to N), $Nim(i)$, can then be defined as the summation of the weighted nimbleness value of each item. Here, the constant values 5 and 0.5 are used to normalize the final value of $Nim(i)$ such that it will fall into the range between 0 to 1:

$$Nim(i)_{0-1} = \frac{\sum_{j=1}^N W(i, j)}{N * 5} + 0.5. \tag{6}$$

Table 1 shows the raw data sets that indicate the response times of 10 test-takers for 10 test items. In this table, the response times for the first test item are shown in the third row. Since $Tmax(1)$ is 10 and $Tmin(1)$ is 1 in this case, the weighted value of the first item is 1.8, and the average response time is 8.2.

Table 1. Example testing data set.

		Student No.													
		$i = 1$	2	3	4	5	6	7	8	9	10	$Tmax(j)$	$Tmin(j)$	$Wun(j)$	$Wavg(i)$
Response time of item j	$j = 1$	1	1	10	10	10	10	10	10	10	10	10	1	1.8	8.2
	2	0.5	2	8	8	8	8	8	8	8	8	8	0.5	1.5	6.7
	3	1	2	9	3	6	2	5	8	12	2	12	1	2.2	5.0
	4	1	2	8	6	2	5	4	6	1	3	8	1	1.4	3.8
	5	1	2	7	2	1	9	1	5	2	6	9	1	1.6	3.6
	6	1	2	8	1	4	8	1	1	6	4	8	1	1.4	3.6
	7	1	2	9	9	9	7	10	2	5	1	10	1	1.8	5.5
	8	2	2	8	12	1	5	11	2	4	2	12	1	2.2	4.9
	9	3	2	5	1	2	4	2	3	2	2	5	1	0.8	2.6
	10	1	2	9	1	8	1	3	5	5	5	9	1	1.6	4.0

The $Nim(i)$ of test-taker 1 (shaded area) in Table 2 is formulated below. For $W(1, 1)$, the weighted nimbleness value of test-taker 1 for item 1 is computed as follows:

$$W(1, 1) = (8.2 - 1)/1.8 = 4, \quad (7)$$

where $W(1, 2) \dots W(1, 10)$ can be computed in the same way. Finally, the nimbleness of thinking of test-taker 1, $Nim(1)$, is formulated as

$$Nim(1) = (W(1, 1) + W(1, 2) + \dots + W(1, 10))/(10 \cdot 5) + 0.5 = 0.912. \quad (8)$$

The nimbleness of thinking of each test-taker can be computed accordingly and is shown in Table 2.

Table 2. Nimbleness of thinking values computed based on Table 1.

		Student No.									
		$i = 1$	2	3	4	5	6	7	8	9	10
Item no	$j = 1$	4.0	4.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
	2	4.1	3.1	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9	-0.9
	3	1.8	1.4	-1.8	0.9	-0.5	1.4	0.0	-1.4	-3.2	1.4
	4	2.0	1.3	-3.0	-1.6	1.3	-0.9	-0.1	-1.6	2.0	0.6
	5	1.6	1.0	-2.1	1.0	1.6	-3.4	1.6	-0.9	1.0	-1.5
	6	1.9	1.1	-3.1	1.9	-0.3	-3.1	1.9	1.9	-1.7	-0.3
	7	2.5	1.9	-1.9	-1.9	-1.9	-0.8	-2.5	1.9	0.3	2.5
	8	1.3	1.3	-1.4	-3.2	1.8	0.0	-2.8	1.3	0.4	1.3
	9	-0.5	0.8	-3.0	2.0	0.8	-1.8	0.8	-0.5	0.8	0.8
	10	1.9	1.3	-3.1	1.9	-2.5	1.9	0.6	-0.6	-0.6	-0.6
$Nim(i)$		0.912	0.843	0.071	0.48	0.467	0.327	0.451	0.466	0.44	0.544

3.2 The Test-Item Answering Ability

The difficulty index and discriminative index of a test item are important factors that can be used to evaluate the performance of test-takers. The test-takers who can correctly answer test items with high difficulty indexes and high discriminative indexes are considered to possess better test-item answering ability. In the following, we formulate the test-item answering ability of a test-taker, $Sol(i)$:

$$Wp(j) = (1 - p(j)) * 5, \quad (9)$$

where $p(j)$ is the difficulty index of test item j and $Wp(j)$ is the weighted value of the difficulty index of test item j . The constant value 5 is used to normalize the final value of $Wp(j)$ such that it will fall into the range between 1 to 5:

$$Wd(j) = D(j) * 5, \quad (10)$$

where $D(j)$ is the discriminative index of test item j and $Wd(j)$ is the weighted value of the discriminative index of test item j . The weighted value of the problem solving ability for test item j can be formulated as follows:

$$Ws(j) = Wp(j) + Wd(j). \tag{11}$$

The $Right(i, j)$ vector records the correct response for item j from test-taker i . The test-item answering ability of test-taker i , $Sol(i)$, is formulated as follows:

$$Sol(i) = \frac{\sum_{j=1}^N Ws(j) * Right(i, j)}{N * 5 * 2}. \tag{12}$$

Table 3. The data set for a test example (difficulty index and discrimination index).

		Student No.											
$Right(i, j)$		$i = 1$	2	3	4	5	6	7	8	9	10	$1-P(j)$	$D(j)$
Response of item j : 1 for correct, 0 for wrong	$j = 1$	1	1	1	1	0	0	1	1	1	1	0.8	0.7
	2	0	0	1	1	1	0	1	1	1	1	0.9	0.6
	3	1	0	1	1	1	0	0	1	1	1	0.8	0.5
	4	1	0	1	1	1	1	0	1	1	1	0.7	0.6
	5	1	0	1	1	1	0	0	1	1	1	0.9	0.7
	6	1	0	1	0	0	1	0	1	1	1	0.8	0.6
	7	0	0	1	1	0	1	0	0	0	1	0.7	0.5
	8	0	0	0	0	1	1	0	0	1	1	0.9	0.6
	9	1	0	0	0	0	0	0	1	0	1	0.8	0.5
	10	1	0	0	1	1	0	0	0	0	1	0.7	0.6

Table 3 shows the data set for one test which has ten test items.

The item answering ability for the ten-item test for test-taker i , $Sol(i)$, is computed in the following manner. For this data set example, the weighted value of the difficulty index of the 1st item for test-taker 1 is determined to be 4 by means of Eq. (9):

$$Wp(1) = (1 (P(1)) * 5 = (1 (0.2) * 5 = 4.$$

The weighted value of the discriminative index of the 1st item for test-taker 1, in this example, is determined to be 3.5 from Eq. (8):

$$Wd(1) = D(1) * 5 = 0.7 * 5 = 3.5.$$

Also, the weighted value of the test-item answering ability for test item 1 for test-taker 1 is 7.5 from Eq. (9):

$$Ws(1) = Wp(1) + Wd(1) = 4 + 3.5 = 7.5.$$

The summation of the weighted value of the test-item answering ability for test-taker 1 is determined to be 48.5 from the numerator of Eq. (10). Finally, $Sol(1)$ is computed as follows: $Sol(1) = 48.5/(10 * 5 * 2) = 0.485$. Table 4 shows the test-item solving ability results for the test data set shown in Table 3.

Table 4. The ten students' item solving ability results based on Table 3.

item#	$Wp(1) + wd(1)$	$Wp(2) + wd(2)$	$Wp(3) + wd(3)$	$Wp(4) + wd(4)$	$Wp(5) + wd(5)$	$Wp(6) + wd(6)$	$Wp(7) + wd(7)$	$Wp(8) + wd(8)$	$Wp(9) + wd(9)$	$Wp(10) + wd(10)$
$j = 1$	7.5	7.5	7.5	7.5	0	0	7.5	7.5	7.5	7.5
2	0	0	7.5	7.5	7.5	0	7.5	7.5	7.5	7.5
3	6.5	0	6.5	6.5	6.5	0	0	6.5	6.5	6.5
4	6.5	0	6.5	6.5	6.5	6.5	0	6.5	6.5	6.5
5	8	0	8	8	8	0	0	8	8	8
6	7	0	7	0	0	7	0	7	7	7
7	0	0	6	6	0	6	0	0	0	6
8	0	0	0	0	7.5	7.5	0	0	7.5	7.5
9	6.5	0	0	0	0	0	0	6.5	0	6.5
10	6.5	0	0	6.5	6.5	0	0	0	3	4.4
$Sol(i)$	0.485	0.075	0.49	0.485	0.425	0.27	0.15	0.495	0.535	0.674

4. THE DESIGN AND IMPLEMENTATION OF A DIAGNOSTIC TEST SYSTEM BASED ON THE ENHANCED S-P MODEL

In this section, we present the design and implementation of a diagnostic test system based on the Enhanced S-P model discussed in section 3. The diagnostic test system based on the Enhanced S-P model is integrated into a web-based CBT system as illustrated in Fig. 4. It integrates four different subsystems, namely, the item management subsystem, online test subsystem, diagnosis subsystem, and user management subsystem.

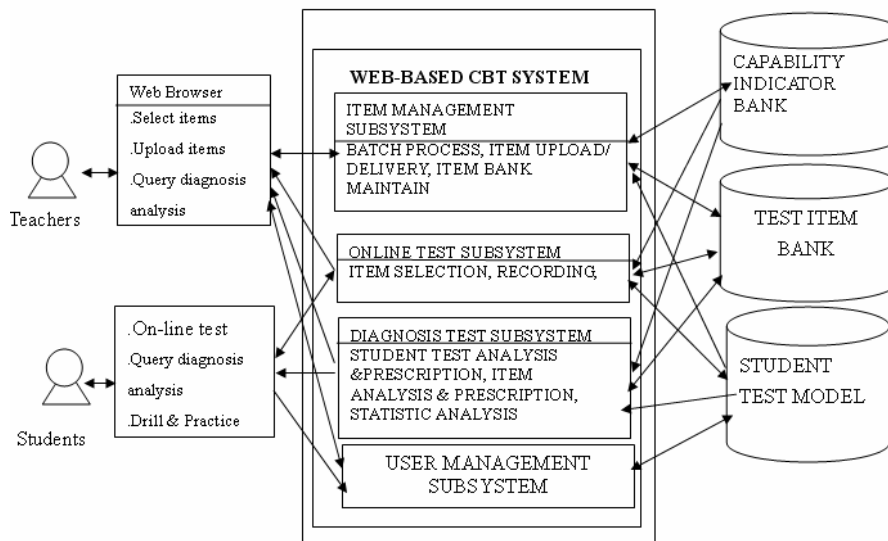


Fig. 4. A CBT system framework.

4.1 The Item Management Subsystem

The item management module is in charge of maintaining both the test item bank and the capability index bank. It controls the delivery of test items to the client side (through wired or wireless devices), and the upload of test item files (from the client side) to the test item bank (the server side) through the http or ftp protocol. It also performs the function of online test sheet composition for test-item creators. This system also supports multimedia test items. Fig. 5 shows the user interface (a single choice type test item with a multimedia presentation) as it might appear during a test. Many different types of test items are supported in this system, including single choice, multiple choice, true-false, singular fill-in-the-blank, multiple fill-in-the-blanks, short answer, question-answer matching, and so on. These different types of test items all have rich presentation formats for both test questions and test answers. For example, one can use a video for the test question and expect the test-taker to submit an audio answer with sound, or similarly, one can use images with a text presentation, sound with images, and so on.

The online test sheet construction method supported in the system allows test sheet creators to create test sheets to meet certain curriculum goals; that is, the test items selected by the teacher comprise the item metadata, and the system uses the defined metadata to monitor the testing process and evaluate students online. Thus, to create a test item sheet, the instructor is required to fill in some parameters associated with the test plan and to select the test layout including the test item type, the test item number, and the grade weighting for the test item, and to select a set of test items from the associated test item bank.

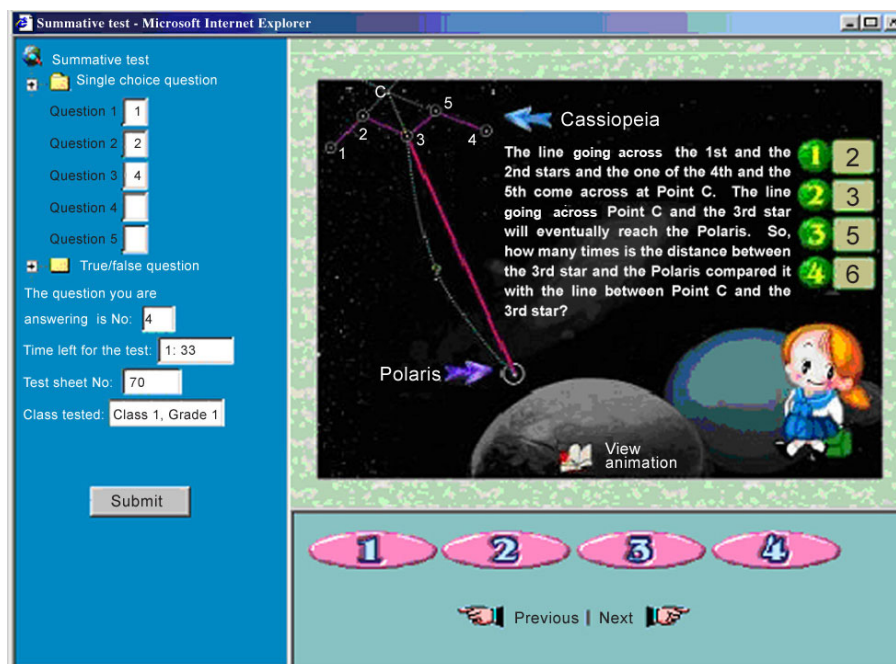


Fig. 5. Example of a multimedia test item.

4.2 The Online Test Subsystem

The online test subsystem, as shown in Fig. 5 supports many features, such as different types of test item selection functions, test item number selection functions, a testing trace mechanism, and test item response uploading. The test item selection approaches supported in this subsystem include automatic random selection and teacher-defined approaches. When a test is activated, this subsystem allows students to take the test online and records the test-taker's response time and other parameters useful for the diagnosis test subsystem. Once the test is finished, this subsystem generates feedback and analysis results, such as how many test items were or were not answered, and how much time was spent on the test by a test-taker.

4.3 The Diagnostic Test Subsystem

As stated in the previous section, the enhanced S-P model was used to design and implement the diagnostic test subsystem. The Entity Relationship (ER) diagram and the data flow of the diagnosis module are depicted in Figs. 6 and 7, respectively.

Fig. 6 depicts the relationship between the online testing subsystem and diagnostic test subsystem. The data collected from the online test subsystem is analyzed in the diagnostic test subsystem. The diagnostic test subsystem first computes the evaluation indexes, such as the student caution index, test item caution index, nimbleness of thinking index, item solving ability index, and so on, based on the data collected from the online test subsystem. Since the original S-P chart is difficult for most users to read, our system allows users to define some common terms based on the correlation with these various indexes obtained through diagnostic analysis. For example, the item prescription needed by teachers to evaluate all of the items and to verify the items' problems includes the disparity index, item caution indexes, discrimination indexes, difficulty indexes, reliability index, and text-based diagnostic description. The individual learner prescription includes the student caution index, item solving ability index, nimbleness index, understood and not understood concepts, a hybrid text-based description of the individual's performance, and suggestions for future learning. The instructional prescription can be used by teachers to improve their pedagogical methods, and it contains the frequency statistics concerning weak performance and suggestions for remedial instruction.

Fig. 7 depicts the Data Flow Diagram (DFD) for diagnosis and statistic analysis in this subsystem.

Basically, the system can be used to diagnose both test items and students with the assistance of the incorporated enhanced S-P model. Through the computation of the various indexes discussed earlier, the system is able to generate 1) some learning advice (prescriptions) for students and parents, and 2) instructional advice (prescriptions) for teachers or instructors.

4.4 The User Management Subsystem

In this subsystem, legal users' accounts, passwords, and useful management information are maintained using the supported functionalities.

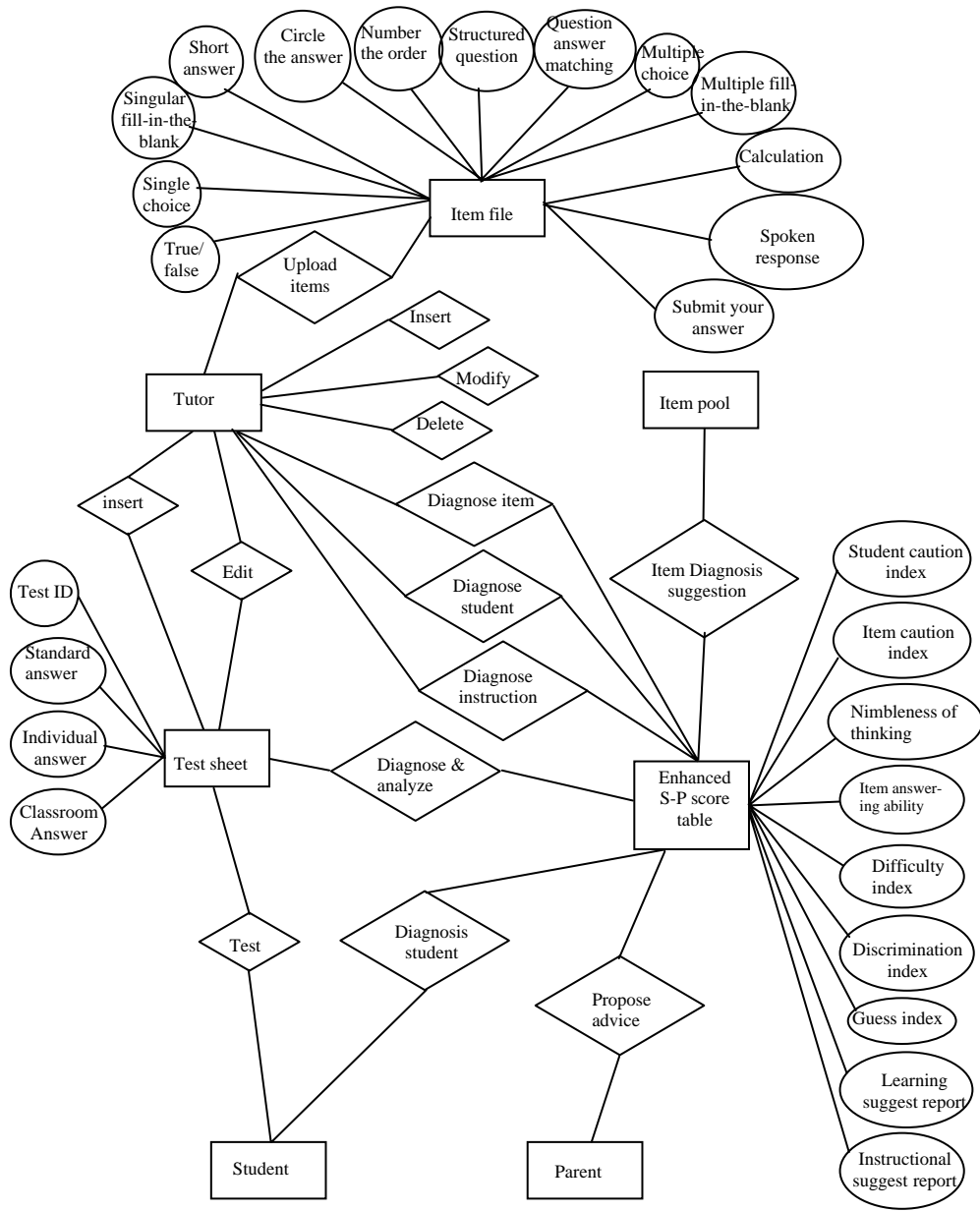


Fig. 6. The ER diagram of the diagnostic module.

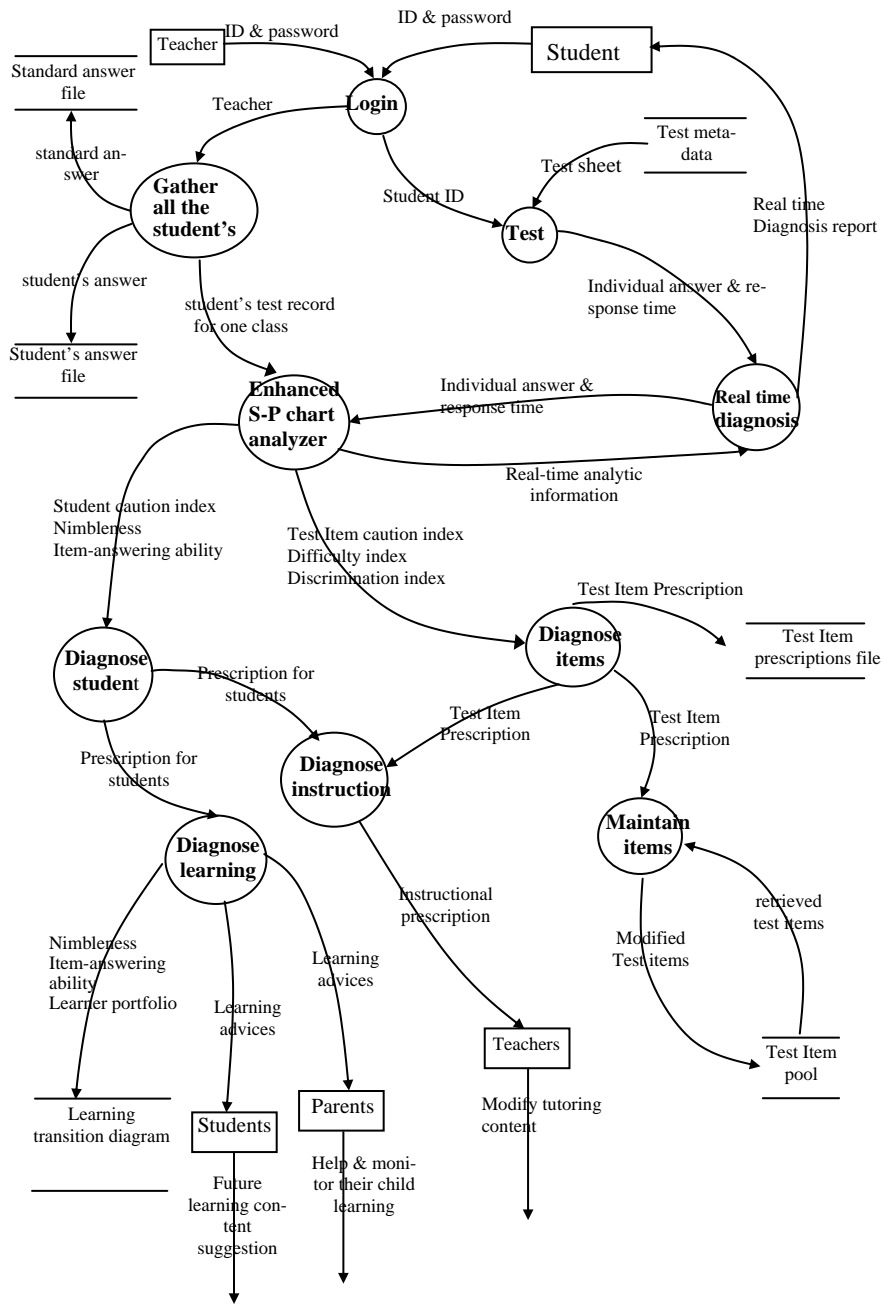


Fig. 7. The DFD of the diagnosis process.

5. APPLICATION AND EXAMPLES

The integrated CBT system has been applied in research conducted at an elementary school and a senior high school. In the following, we will present some snapshots of the user interface of the proposed CBT system, which was used to collect data in the above-mentioned study.

5.1 Test Item Uploading

A snapshot of the test item uploading module is shown in Fig. 8. The test item uploader needs to select the type of test item first. Fig. 8 indicates that there are 14 possible types of test items.

Test-items Upload	
Single choice type	True/False type
Singular-fill-in-the-blank type	Calculation type
Short answer type	Voice-your-answer type
Submit-your-answer type	Question-answer matching type
Multiple answers in multiple choice type	Multiple-fill-in-the-blanks type
Structured Question type	Circle the answer type
Numbering-the-order type	Mixed type

Fig. 8. Fourteen types of test items supported in the system.

Once a type of test item has been selected, the system asks the test item uploader to fill in the required parameters associated with the test item shown in Fig. 9. These parameters include the subject, item type, test purpose, capability index, and the name of the test.

Question Upload	
Subject	Question uploaded to : Science and technology domain/Book 1/Chapter 1
Question	<input type="text"/>
Answer	<input type="radio"/> A <input type="radio"/> B <input type="radio"/> C <input type="radio"/> D Sharing? <input checked="" type="checkbox"/>
Question type	<input type="radio"/> Single choice type
Test purpose	<input type="radio"/> Placement test <input checked="" type="radio"/> Formative test <input type="radio"/> Summative test
Competence indicators	<input type="text"/>
Target concept	<input type="text"/>
File upload interface	<input type="text"/> <input type="button" value="View..."/> <small>(File upload path)</small>
Test beginning page	<input type="text"/> <input type="button" value="Upload"/> <small>(for example, index.htm)</small>

Fig. 9. Interface dialogue for uploading a test item.

When the test item uploader fills in the capability index parameters, the system automatically shows the interface for selecting the capability index for the test item as shown in Fig. 10. The uploader is required to select an associated (predefined) capability index for the test item. This capability index table is based on the Ministry of Education's Taiwan 9-year mandatory basic education program (grade 1-9 curriculum).

Domain : Science and technology, Domain stage No. : 1

Competency index :

Choice	Learning Areas	Competence Indicators	Learning stage	Concept name	Indicator description
<input type="radio"/>	Science and technology	1-1-1-1	1	Computer network concept	can understand the categories of protocol of computer network
<input type="radio"/>	Science and technology	1-1-1-2	1	Computer network application	can understand the application of computer network
<input type="radio"/>	Science and technology	1-1-2-1	1	Graphics format	can understand the specification and kinds of computer graphics format

Page NO : 1/20, total records: 60

Fig. 10. Interface for selecting the capability index for the item file.

5.2 Online Test Sheet Composition and Management

In the previous section, we explained how a test item is uploaded to the CBT system. In this section, we will explain how a test sheet is composed. The test sheet creator needs to use the test item management system for test sheet composition. Fig. 11 depicts the user interface for the test sheet information that the test sheet creator needs to provide.

Author of the test	Teacher name
Semester #	Year: 2002, Semester : 2nd
Title of the test	Subject: <input type="text" value="BOC"/> Exam Type: <input type="text" value="1st mid term"/>
Class tested/ test time	Class: <input type="text" value="class 1"/> Answering time: <input type="text" value="50"/> min
Test sheet description	<input type="text"/>
<input type="button" value="Submit test metadata"/>	

Fig. 11. User interface for test sheet creation.

The information that needs to be specified includes the subject of the test, the class name (or grade level), the maximum test time allowed for the test, and a short statement describing the goal of the test. After the above-mentioned basic information is specified and submitted to the server, the test creator is asked by the system to set up the layout and composition of the test sheet as shown in Fig. 12.

The Test Sheet Plan and Layout

Total questions		Total score	
24		100	
Question type	Number of questions of each type	Score for each question	
<input checked="" type="checkbox"/> Single choice type	Number: 10	Score: 4	
<input checked="" type="checkbox"/> True/False type	Number: 5	Score: 4	
<input checked="" type="checkbox"/> Single-fill-in-the-blank type	Number: 5	Score: 4	
<input checked="" type="checkbox"/> Calculation type	Number: 4	Score: 5	
<input type="checkbox"/> Short answer type	Number: 1	Score: 1	
<input type="checkbox"/> Speak-your-voice type	Number: 1	Score: 1	
<input type="checkbox"/> Submit -your-answer- type	Number: 1	Score: 1	
<input type="checkbox"/> Question-answer matching type	Number: 1	Score: 1	
<input type="checkbox"/> Multiple answers in multiple choice type	Number: 1	Score: 1	
<input type="checkbox"/> Multiple-fill-in-the-blanks type	Number: 1	Score: 1	
<input type="checkbox"/> Structured question type	Number: 1	Score: 1	
<input type="checkbox"/> Circle-the-answer type	Number: 1	Score: 1	
<input type="checkbox"/> Number-the-order type	Number: 1	Score: 1	

Submit test layout

Fig. 12. Interface for the test sheet plan and layout.

Current database: Science and technology domain/ Book 1 > The following is 25 items for this sheet.

Available questions	Question No.	Printable	Correct answer count	Times being selected	Checkbox for selecting each item
When the data is transmitted in the computer network, which method is the most efficient? (A) Simplex wire (B) Half duplex (C) Full duplex (D) Multiple channel.	225	Yes	0	0	<input checked="" type="checkbox"/>
Which protocol is used for the service of Email-sending in the Internet? (A)POP3 (B) SMTP (C) NetBIOS (D) TCP/IP (E) IPX/SPX	226	Yes	0	0	<input checked="" type="checkbox"/>
Which of the following isn't the service provided in the Internet? (A) WWW (B) E-mail (C) VCD (D) BBS (E) Netnews.	227	Yes	0	0	<input checked="" type="checkbox"/>
If we will find the data in an array that contains 100 keys, how many comparisons must be made in the best case of linear search? (A) 0 (B) 1 (C) 50 (D) 51 (E) 100.	228	Yes	0	0	<input checked="" type="checkbox"/>
Suppose four bits are used to represent a graphic pixel, how many colors can be presented in that computer? (A)4 (B)8 (C)16(D)256 (E)65536.	229	Yes	0	0	<input type="checkbox"/>
5 points per item, Number of items are 10.	You have selected: 4 items	Total grade per unit: 20 points	Single choice type		
5 points per item, Number of items are 10.	You have selected: 5 items	Total grade per unit: 25 points	True / False type		
Submit		Return to previous page			

Fig. 13. Interface for selecting of items.

As indicated in Fig. 12, the test creator has set up the test sheet with 10 single choice test items, 5 true/false test items, 5 single fill-in-the-blank test items, and 4 calculation type test items, with each test item of the calculation type receiving a weighting of 5% of the total test score, and each test item in the rest of the type of test item receiving a weighting of 4% of the total test score. The test sheet's layout can now be created. To do so, the test sheet creator needs to select test items from the test-item bank according to the layout just created. Fig. 13 shows how test items are retrieved according to the associated capability index.

As shown in Fig. 13, friendly user interfaces are implemented in the CBT system to assist test sheet creators. Once the test sheet is created and completed, the creator clicks the submit button to deposit the test sheet in the database.

5.3 Online Testing and Diagnosis

To take an online test, the test-taker (student) needs to login to the system with appropriate identification such that the system can identify which set of test sheets he or she can use to take the test as shown in Fig. 14. The test-taker then selects a test to take.

Online Testing

Current Database: Science and technology domain/Book 1

Click the button for test	Test title	Subject	Test sheet description	Author of the test sheet
<input type="button" value="Start to test"/>	Year: 2002, Semester: 2, Formative test	Science and technology	Computer network and program concept	Arthur
<input type="button" value="Start to test"/>	Year :2002, semester:2 ,2nd term test	Science and technology	Star and solar system	Arthur

Fig. 14. Online test menu.

After all of the online test-takers have finished filling in their predefined test sheets and have submitted them to server, the system begins to diagnose the test sheets. At this point, the diagnosis system is launched. The diagnosis system uses the enhanced S-P model discussed earlier to compute various indexes associated with the test item, test sheet, and test-takers, and then interprets these indexes by producing item prescriptions and instructional advice for test-takers and instructors, respectively, as illustrated in Figs. 15 and 16.

Prescription for item				
Test title	Test sheet No:69, for year 2002, 2nd semester, formative test of ECC, Grade 1 class 1			
Test sheet description	Computer network and program concept			
Test type	Single-choice			
Question #	4Th item			
Hyperlink for item previewed	Click here for previewing items			
Average answer time	0min 23 seconds			
Analysis of the S-P score	Item caution index	0.09	Difficulty index	0.55
	Discrimination index	0.9	Item category	B
General suggestions	Item caution index : The item caution index is moderate, no correction needed. Difficulty index : The item difficulty is moderate, but works better on average students. Discrimination index : Student's scores can fully reveal their leaning status.			

Fig. 15. Example of an item prescription after queried by the teacher.

Fig. 15 shows that the caution index, the difficulty index, the discrimination index, and the average response time for the 4th test item are 0.09, 0.55, 0.9, and 23 seconds, respectively. The diagnosis description shows that the item is perfect, and that it does not need to be modified. Also, the difficulty level is medium and suitable for average students. Fig. 16 shows that 7, 2, and 19 students were unable to achieve a capability index of 1-1-1-1 (the concept of network communication), 1-1-1-2 (the application of a network), and 1-1-2-2 (the programming concepts of loop and sort), respectively. The system shows that the instructor should provide remedial coursework on the programming concept for the students in that class.

Instructional advice	
Test title	Test sheet No:69, for year 2002, 2 semester, formative test of BCC, Grade 1 class 1
Test sheet description	Computer network and program concept
classroom	Grade 1 class1
Student numbers	29
Number of students who failed to achieve the required capability index	1-1-1-1(7 students), 1-1-1-2(2 students), 1-1-2-2(9 students)
Number of students who failed to understand the target concept in this teaching session	Computer network concept (7 students), computer network application (2 students), program loop(8 students), sorting algorithm (7 students).
Remedial coursework suggested	Remedial capability index : 1-1-2-2, Remedial learning concept : Program loop

Fig. 16. Example of instructional advice.

Class # Student's name:

Prescription for student				
Test title	Test sheet No:69, for year 2002, 2 semester, formative test of BCC, Grade 1 class 1			
Test sheet description	Computer network and program concept			
Index for the student	Caution index	0	Nimbleness of thinking	0.67
	Item-solving capability	0.23	Learning category	A
Competence indicators achieved	1-1-1-1(5 questions/5 correct) , 1-1-1-2(5 questions/5 correct) , 1-1-2-1(5 question/5 correct).			
Target concepts understood	Computer network concept (4 items, all correct), Computer network application(7 questions, all correct), graphics format(5 item, all correct).			
Competence indicators uncompleted	1-1-2-2 (4 out of 5 answers are incorrect).			
Concepts not understood	Program concept (misconception ratio: 2/3), sorting algorithm (misconception ratio: 2/2).			
Diagnostic suggestion	Caution index: The student is doing well. She understands target concepts quickly and correctly except for the program concept. Nimbleness of thinking: Time spent on each question is almost the same. Item-solving capability: Problem solving capability should be improved, especially for the program concept.			
Future learning suggestion	Next competence indicator: 1-1-2-2 ; Next target concept: Program loop; sorting algorithm.			

Fig. 17. Example of a prescription for an individual learner following system diagnosis

The test-takers also can query their individual learning status and get the suggestions after the test. Fig. 17 shows an example of an individual learner's prescription. The system gives the diagnosis description for these indexes. In this example, the student caution index, the nimbleness of thinking index, and the test-item answering ability index are 0, 0.67, and 0.23, respectively. The test-taker in this case has attained the curriculum

goal with capability indexes of 1-1-1-1 (network communication), 1-1-1-2 (network application), and 1-1-2-4 (the format of the picture file). However, test-taker's understanding of the programming concept is very poor, as indicated by the capability index of 1-1-2-2, the student should accept the remedial course in this subject area.

Table 5. Example of an SP score table for students in one senior high school.

Class#:101		Item No.																													
ID	name	10	11	15	7	5	14	12	3	16	9	2	1	13	8	4	18	6	17	20	19	score	CR%	CS	LC	NT	ISA				
1	Stu01	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20	1	0	A	0.71	0.33				
10	Stu10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20	1	0	A	0.62	0.33				
12	Stu12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	19	0.95	0	A	0.7	0.3				
32	Stu32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	19	0.95	0.14	A	0.62	0.3				
2	Stu02	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	19	0.95	0.82	A'	0.43	0.3				
17	Stu17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	18	0.9	0.07	A	0.63	0.28			
3	Stu03	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	17	0.85	0.67	A'	0.54	0.26				
29	Stu29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	16	0.8	0	A	0.64	0.23				
30	Stu30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	16	0.8	0	A	0.6	0.23				
24	Chang	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	16	0.8	0	A	0.67	0.23				
21	Stu21	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	14	0.7	0.06	B	0.52	0.17				
9	Stu09	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0	0	0	0	14	0.7	0.29	B	0.45	0.18				
28	Stu28	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	13	0.65	0	B	0.43	0.14				
14	Stu14	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	13	0.65	0	B	0.45	0.14				
31	Stu31	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	13	0.65	0.02	B	0.31	0.14				
25	Stu25	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	13	0.65	0.02	B	0.25	0.14				
26	Stu26	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	13	0.65	0.08	B	0.19	0.13				
23	Stu23	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	0	1	0	0	0	13	0.65	0.36	B	0.15	0.18				
11	Stu11	1	1	1	1	1	0	1	1	1	1	1	0	0	1	0	1	0	0	0	1	13	0.65	0.42	B	0.3	0.16				
8	Stu08	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	0	0	1	0	0	13	0.65	0.44	B	0.45	0.17				
4	Stu04	1	1	1	1	1	1	0	1	1	1	0	0	0	1	0	1	0	1	1	0	13	0.65	0.62	B'	0.57	0.15				
27	Stu27	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	12	0.6	0	B	0.5	0.11				
18	Stu18	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	12	0.6	0	B	0.48	0.11				
16	Stu16	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	12	0.6	0	B	0.51	0.11				
15	Stu15	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	12	0.6	0.09	B	0.49	0.11				
22	Stu22	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	1	0	0	0	12	0.6	0.16	B	0.58	0.13				
13	Stu13	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	11	0.55	0.14	B	0.59	0.09				
6	Stu06	1	1	1	1	0	1	1	1	0	0	0	0	1	0	1	0	0	1	1	0	11	0.55	0.78	B'	0.44	0.15				
7	Stu07	1	1	1	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	7	0.35	0.46	C	0.63	0.09				
Correct responses		29	29	29	28	28	27	27	26	26	26	23	19	18	17	14	13	13	8	8	6	414									
Correct %		1	1	1	0.97	0.97	0.93	0.93	0.9	0.9	0.9	0.79	0.66	0.62	0.59	0.48	0.45	0.45	0.28	0.28	0.21	Note									
Item caution index		0	0	0	0	0.56	0.19	0.19	0.29	0.22	0.44	0.39	0.14	0.52	0.36	0.09	0.37	0.06	0.47	0.3	0.2	(1)CR%: Correct Response % of the student									
Diff index		0.95	0.95	0.95	0.9	0.9	0.9	0.85	0.9	0.85	0.8	0.75	0.65	0.6	0.55	0.55	0.5	0.55	0.35	0.4	0.25	(2)CS: Caution index of the Students									
Disc index		0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.2	0.3	0.2	0.5	0.7	0.6	0.7	0.9	0.4	0.9	0.3	0.4	0.5	(3)LC: Learning Category									
Item category		A	A	A	A	A'	A	A	A	A	A	A	A	A'	A	B	B	B	B	B	B	(4)NT: the index of Nimbleness of the Thinking									
KR-20 Coefficient		0.773										Disparity Coefficient (D*)										1.367					(5)ISA: Item-Solving Ability				

5.4 The S-P Score Table and Statistic Diagrams

Table 5 shows an example of an enhanced S-P score table obtained from the diagnostic test on the 1st year of senior high school in Taiwan (which is equal to the 10th

grade in the U.S. system). This enhanced S-P table provides the entire diagnostic analysis results for the test-takers and test items. The bottom part of the table shows the correct response percentage, the caution index, the difficulty index, and the discrimination index for each test item. The caution indexes for the 5th test item and the 13th test item are both above the normal value, indicating that these two test items should be redesigned or modified. The discrimination indexes of test items 10, 11, 15, 7, 5, 14, 3, and 9 are lower than the normal value, so these test items cannot differentiate the learning performance of the test-takers. The KR-20 coefficient is 0.773, which shows that the test sheet is stable and reliable. The disparity index for this test is 1.367, which is over 0.6, indicating that some of the test items should be improved and modified. The scatter diagram of the test items' caution index and the bar chart of item category in one test are illustrated in Fig. 18.

As for student diagnosis, the right side of the Table 5 shows the correct response percentage, the caution index, the nimbleness of thinking index, and the item-answering ability index for students. The caution indexes of students' number 2, 3, 4, and 6 are all within the normal range. This indicates that these students may have been careless during the test or lacked understanding of the learning contents or unstable during his study. The scatter diagram of students' caution index and the bar chart of students' learning categories in one test are illustrated in Fig. 19.

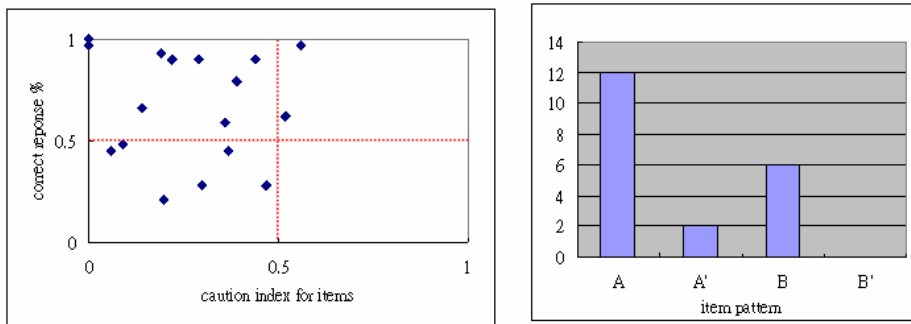


Fig. 18. Scatter diagram of items' caution index and bar chart of item category in one test.

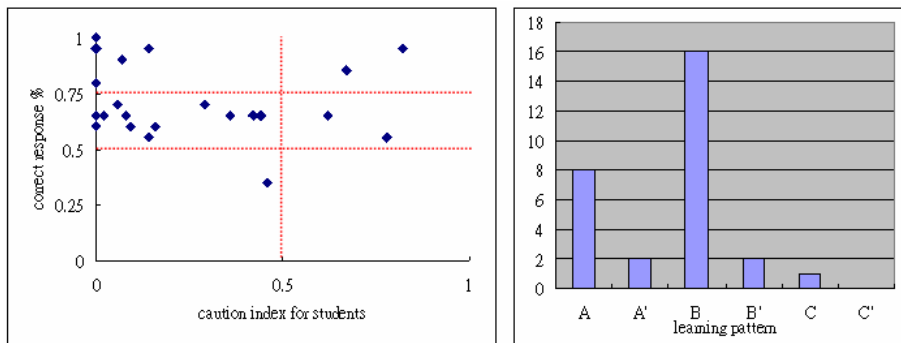


Fig. 19. Scatter diagram of students' caution index and bar chart of students' learning category in one test.

Table 5 also shows that the nimbleness index can be used to tell which students responded quickly to certain test items. A student with higher nimbleness index possesses better mastery of the learning content. For instance, students number 10, 12, and 32 have the same number of correct responses in the test, but their nimbleness indexes are different. In this case, it is appropriate to conclude that student number 12 has the better mastery of the learning content.

In traditional test assessment, all students who have the same number of correct responses get the same score or ranking. In this system, we can further differentiate among students who have the same score or ranking based on their item-answering ability. In this example, both students number 21 and 9 have 14 correct responses (or 70%) in the test, but their item-answering ability is 1.7 and 1.8, respectively.

6. CONCLUSIONS

The diagnostic test system plays an important role in distance learning and web-based instruction because of its functions for assessing learner performance and diagnosing learning problems. As a result, a S-P chart was proposed for this purpose. However, the current S-P model cannot take the test-taker's response time into consideration. In this study, the nimbleness of thinking index has been proposed to enhance the S-P model. It possesses time dependent information and can be used to understand the response speed of test-takers. We have also proposed the test-item answering ability index for analyzing an examinee's test-item answering capability in a test. In addition, a proposed diagnostic test system based on the Enhanced S-P model has been designed and implemented. This system has been integrated into a web-based CBT system, which has four different subsystems, namely, the item management subsystem, online test subsystem, diagnosis test subsystem, and users management subsystem.

This integrated CBT system has been applied in a study conducted at an elementary school and a senior high school. Application examples have been given to demonstrate the applicability of this system. Various useful charts and prescriptions can be generated by the system and used by instructors to diagnose their instructional approaches and by students to diagnose their learning performance. This information can provide guidelines to help instructors improve their pedagogy and supply course remedy advice to students after the test.

ACKNOWLEDGMENTS

This research is supported in part by the Bestwise International Computing Co., CAISER (National Chiao Tung University, Taiwan), and National Science Council (Taiwan).

REFERENCES

1. A. S. McDonald, "The impact of individual differences on the equivalence of computer-based and paper-and-pencil educational assessments," *Computers and Educa-*

- tion, Vol. 39, 2002, pp. 299-312.
2. C. Chien, "Constructing a computer-assisted testing and evaluation system on the world wide web – the CATES experience," *IEEE Transactions on Education*, Vol. 43, 2000, pp. 266-272.
 3. D. Harnisch and R. L. Linn, "Analysis response patterns: questionable test data and dissimilar curriculum practices," *Journal of Educational Measurement*, Vol. 18, 1981, pp. 133-146.
 4. D. M. Switzer and M. L. Connell, "Practical applications of student response analysis," *Measurement in Education*, the annual meeting of the national council of 1989, pp. 53-65.
 5. D. W. Johnson and R. T. Johnson, *Meaningful Assessment*, Allyn and Bacon, 2002.
 6. H. Y. Wu, "Software based on S-P chart analysis and its applications," *National Science Council*, NSC-86-2511-S-252-001-CL, 1998.
 7. I. Chacko, "S-P chart and instructional decisions in the classroom," *International Journal of Mathematical Education in Science and Technology*, Vol. 29, 1998, pp. 445-450.
 8. J. Salvia and J. E. Ysseldyke, *Assessment*, 8th ed., Houghton Mifflin Company, 2001.
 9. M. N. Yu, *Educational Test and Assessment: Performance Test and Instructional Assessment*, Psychological Publishing Co., Ltd., 1987. (in Chinese)
 10. S. L. Blixt and T. E. Dinero, "An initial look at the validity of diagnoses based on Sato's caution index," *Educational and Psychological Measurement*, Vol. 45, 1985, pp. 55-61.
 11. T. E. Dinero and S. L. Blixt, "Information about tests from Sato's S-P chart," *College Teaching*, Vol. 36, 1988, pp. 123-128.
 12. T. Sato, *The Construction and Interpretation of the S-P Table – Instructional Analysis and Learning Diagnosis*, Tokyo, Japan, Meiji Tosho, 1975. (in Japanese)
 13. T. Sato, *The S-P Chart and the Caution Index*, NEC Educational Information Bulletin 80-1, C&C Systems Research Laboratories, Nippon Electric Co., Ltd., Tokyo, Japan, 1980. (in Japanese)
 14. T. Sato, *Introduction to S-P Curve Theory Analysis and Evaluation*, Tokyo, Japan, Meiji Tosho, 1985. (in Japanese)



Deng-Jyi Chen (陳登吉) received the B.S. degree in Computer Science from Missouri State University (Cape Girardeau), U.S.A., and the M.S. and Ph.D. degrees in Computer Science from the University of Texas (Arlington), U.S.A. in 1983, 1985, and 1988, respectively. He is now a Professor in the Computer Science and Information Engineering Department of National Chiao Tung University (Hsinchu, Taiwan). Prior to joining the faculty of National Chiao Tung University, he was with National Cheng Kung University (Tainan, Taiwan). So far, he has published more than 130 related papers in the areas of software engineering (software reuse, object-oriented systems, and visual requirement representation), multimedia application systems (visual authoring tools), e-learning and e-testing

systems, performance and reliability modeling, the evaluation of distributed systems, and computer networks. Some of his research results have been transferred to the industrial sector and used in product designs. So far, he has been a chief project leader for more than 10 commercial products. Some of these products are widely used around the world. He has received both research and teaching awards from various organizations in Taiwan and has served as a committee member for several academic and industrial organizations.



Ah-Fur Lai (賴阿福) is currently a doctoral candidate in the Department of Computer Science and Information Engineering at National Chiao Tung University. He is now an Associate Professor in the Department of Mathematics and Computer Science Education at Taipei Municipal University of Education. His research interests include e-learning, and computer-based testing and assessment.



I-Chang Liu (劉奕樟) received the M.S. degree in e-Learning from National Chiao Tung University in 2003. He is currently the director of academic affairs at National Chung-Li Home Economics and Commercial Vocational High School.