Developing a Web-Based System for Testing Students' Physics Misconceptions (WEBSYSTEM) and its Implementation

Minkee Kim* · **Jaehyeok Choi**[1] · **Jinwoong Song**

Seoul National University · Korea Education & Research Information Service[1]

Abstract: Several studies have attempted to test students' misconceptions of physics and to provide teaching strategies in order to repair them. The results from these studies have revealed that the diagnosis of students' misconception is crucial, although they often failed to grasp the practice of its implementation. In terms of being a type of methodology for science education, the Internet allows large-scale surveys and investigations to be carried out in a relatively short period of time. This paper reports the results of the development, implementation, and evaluation of a WEb-based SYStem for TEsting students' Misconceptions in physics (WEBSYSTEM) aimed at three groups (science educational researchers who study students' physics conceptions using the system as a detector, school science teachers who practice it as an instructional material, and students who benefit from it for their self-directed learning). The web-based testing system is based on a review of the instructional development strategies of ADDIE (Gustafson, Branch, 2002; Rha, Chung, 2001). Results showed that WEBSYSTEM could work effectively as a multi-purposed tool for the three target groups with a further partial revision, providing educational researchers with resourceful data to study students' misconceptions in physics. Issues of administrative strategies, reexamination of questionnaires, and international collaboration via WEBSYSTEM are discussed.

Key words: physics misconception, WEBSYSTEM, online evaluation, computer-based evaluation, research methodology

I. Introduction

A number of previous studies have disclosed that students have various misconceptions in physics. Many different strategies have been suggested to change students' misconceptions (diSessa, 2002; Vosniadou, 2002; Chi *et al.*, 1994; Posner *et al.*, 1982; Kim, 1991; Oh, 1998; Park *et al.*, 1994). One of their crucial findings is that we need to have detailed knowledge concerning students' conceptions to apply these strategies (Taber, 2003). Individual physics teachers may not be capable of investigating these strategies due to lack of time and resources to accumulate information and analyze the data from students. As an aid in teaching and learning material, Fensham (2004) reported that a well-structured conception inspection tool from science educational studies could be employed by science teachers to identify students' various science conceptions and by students to glance at others' ideas of specific science contents to generate discussion. According to Fensham's proposed practicable usage of the research methodologies in science classes, science misconception researchers should develop their tools to the point where the three target groups (science educational researchers, science teachers, and students) benefit as a whole, proving the practice of their own research.

The results from a large-scale web-based survey in Korea, involving 1,670,000 students (22% of the total number of students in Korea) from 5th to 12th graders, revealed many advantages of using the Internet such as: saving time, transferring responses into a database with fewer systematic errors, avoiding the loss of data, and the ability to cover a wider geographic area (Korea Science Foundation, 2003). This experimental attempt at creating a new research methodology has shown that utilizing a web-based approach is a feasible means of gathering survey data (Mertler, 2003). Furthermore, in later studies, the Internet was proven to be an educational medi-

*Corresponding author: Minkee Kim (amour03@snu.ac.kr)
**Received on 23 March 2006, Accepted on 6 February 2007

um that students generally prefer to other conventional media (Ministry of Education & Human Resources Development, 2003; Engelbrecht, Harding, 2004), though Wade & Lyng (2000) pointed out the issue that the automated evaluation systems via the Internet need to be validated with various case studies and also be implemented in diverse contexts.

Research concerning students' conceptions of school science has been achieved through the effective methodology of the Internet. Firstly, students communicated more frequently and effectively with the Internet, generating students' conceptual changes (Oshima *et al.,* 2004; Shim *et al.,* 2005; Tao, Gunstone, 1999). Students tended to engage more in Internet-based communication without obstacles of time or space. Secondly, the Internet -based science simulations could become among the foremost educational materials for deriving conceptual change in students (Park *et al.,* 2003; Zietsman, Hewson, 1986). These simulations turned out to be the reliable tools for visualizing science concepts, enhancing interest among students, helping students to develop laboratory skills, and encouraging creative learning. Lastly, computer aided learning tools were introduced to build structural representations of the knowledge from school science (Zele, Lenaerts, 2004). The researchers developed a computer-aided concept-mapping tool, QMap 0.4 (only available in MS Windows) to provide an electric worksheet as a meta-cognitive presentation tool for students and to automate research data of students' science conceptions of science educational researchers.

Even though the Internet-based tools adopted in the field of science education were proven to be teaching and learning materials or inevitable research methodologies, the detailed procedure of developing them has not been elucidated, which demands cross -field work with the educational technology. Little literature exists clarifying specific developments to assist science educational researchers in reproducing and improving previously authenticated tools in the pursuit of further research. Thus, the aim of this study is to develop a web-based system for testing students' physics misconceptions (WEBSYSTEM) that is beneficial to three groups (science educational researchers, whose main studies are students' physics conception using the system as a detector; school science teachers who practice this system as an instructional material; and students who benefit from this system for their self-directed learning), including implementation as well as evaluation with detailed developmental procedures.

II. Literature Review

Andriole (1997) presented the requirements-driven methodology as his early model of instructional development. With this methodology, developers first derive detailed requirements from target users or administrators. Their findings of the first procedure consist of purposeful requirements and functional requirements. Purposeful requirements analysis requests the reasons why such systems need to exist and how their objectives may be achieved. Meanwhile, functional requirements include the specific functions the target users or administrators expect the system to conduct. The derived requirements consequently provide guidelines to series of following procedures (design, development, delivery and evaluation) throughout the development process.

Research on educational technology has succeeded in creating a developmental model that commences with requirements analysis, and has suggested a revised model for developing web-based instructional system. Gustafson & Branch (2002) proposed a concrete definition of instruction development model that consists of five procedures: (1) analysis of setting and learners' needs; (2) design of a set of specifications for a learner environment; (3) developing parts of the system for learners and management; (4) implementing the resulting system; and (5) conducting formative or summative evaluations of the system. These five procedures have been referred to as ADDIE (Analysis, Design, Development, Implementation, and Evaluation). In the Korean school context, its genetic procedures for developing web -based instructional system have also been reported as a convincing instructional model (Rha, Chung, 2001).

The authors in this research adopted the ADDIE model to provide specific information during the development of WEBSYSTEM by documenting products from each procedure and by defining sets of relationships among them.

III. Method & Procedure

This chapter introduces what we have planned for perusing this study to be an informative developmental study. With five procedures from ADDIE, documentations are subcategorized and presented respectively here and at the following chapter.

1. Requirement Analysis

In order to clarify the aims of developing WEBSYSTEM and to explore possible forms of system implementation, the authors conducted research meetings: seminars with educational researchers, teacher interviews and a student survey, and reviewed literature in the field of students' science conception. These efforts were used to gather the requirements for a web-based system to test students' physics misconceptions (WEBSYSTEM). Research meetings and seminars with educational researchers were held throughout the twelve-month development period, while teacher interviews and a student survey were conducted in the succeeding procedure at a school located in Seoul.

2. System Design

Functions were designed and categorized for WEBSYSTEM based on items from the 'Requirement Analysis'. Literature was reviewed for strategies in computer/Internet-based assessment development. According to the literature, supplemental training should be concerned with eliminating the practical difficulties in a composition test with items presented in computer-print form (Russell, Tao, 2004). The differences between the paper & pencil format test and the web-format test (Mertler, 2003), and administrational issues of computer-based assessment (such as back-up machines, validity of testing materials, delivery of materials, faculties training, and scheduling) were also included into the 'System Design' (Rabinowitz, Brandt, 2001).

3. System Development

One of the three authors programmed WEBSYSTEM with PHP (a web-based program language) and Mysql (a database program). Web-based programs for assessment, community and graphic reports were reviewed and applied into each function of the 'System Design', while other functions were programmed by the author himself.

4. System Implementation

Instructional system designers have used qualitative methods for surveys, interviews and observations during the primary analysis and evaluation phases (e. g. Savenye, Robinson, 2003). Qualitative approaches tend to be more holistic and process -oriented on an operational level (Gay, 1996). In order to capture the richness and the personal perspectives of WEBSYSTEM with the qualitative approaches, the authors gathered data from teacher interviews, a student survey and teacher's after-class reports, as well as accumulated students' test results into the database. In total, 247 (161 grade 10 students and 86 grade 11 students) students and their physics teacher were involved in this study.

An administrative meeting with the teacher was held in order to confirm the possibilities for students' accessibility to the Internet and to schedule the intervention, both of which are critical factors for computer/Internet-based testing (Mertler, 2003; Rabinowitz, Brandt, 2001). The research methods for the teacher interview, teacher's after-class report and student survey are presented in the following section, III-(5).

Test materials for WEBSYSTEM from 'Wave and Light' misconception test questionnaires from a previous study were equipped into WEBSYSTEM. Out of the forty questions, ten of them were selected from Song *et al.* (2004)'s physics misconception test, examining their relevance into the school curriculum with teacher K who was in charge of the sample students.

5. System Evaluation

A 'System Evaluation' was conducted to evaluate the achievement of items in the 'Requirement

Table 1
Categories and detailed questions of teacher interview

Category	Question
Experience in class	Did you set up any administrative strategies to increase student participation?
	Did you make any documents for the notification?
	Did your students have difficulties in taking the test?
	Did your students have questions about the test?
Analysis of test results	Does students' result from WEBSYSTEM have any bearing on correlation with students' physics achievement?
	Do test results vary between teachers?
	Are test results applicable to your physics class?
Opinion of system	Is the format of the web reports useful?
	What is the most informative data among 3 types of Test Reports?
	Do you think the descriptive responses are useful?
	What would you like to have changed or improved?

Analysis'. The data were documented through teacher interviews, teacher's after-class report and student survey. Teacher interviews included the categories of 'Experience in class', 'Analysis of test results', and 'Opinion of system (See Table 1.). The teacher's after-class report included the three issues of 'Comments on using the WEBSYSTEM in physics classes', 'Examples of using the WEBSYSTEM in physics classes', and 'Risks of using the WEB-SYSTEM in physics classes'. A student survey was conducted with open-ended questions and including the three issues of 'Functional inconvenience', 'Benefits of WEBSYSTEM' and 'Risks of WEB-SYSTEM'.

Data were analyzed through a summative evaluation of each of the items found in the 'Requirement Analysis'. We evaluated each item by grading them as follows: grade A, B, C, and D. Grade A presented grade for items of the 'Requirement Analysis', which had been achieved successfully and that had relevant functions that were designed adequately. Grade B presented items that were achieved partially, though they had relevant functions they were designed adequately. If an item of the 'Requirement Analysis' required additional data processing and analysis, it was graded as Grade B. Grade C represented items of the 'Requirement Analysis', which had proven not been achieved and had relevant functions that needed to be revised. Grade D was used for items that had no relevant functions.

IV. Results and Discusstion

1. Requirement Analysis

The authors reviewed the literature (Taber, 2003; Song *et al.,* 2004) and paid attention to what they suggested for further studies in order to extract items that corresponded to the goals of the WEBSYSTEM. The authors held research meetings eight times and seminars three times with a total of 10 – 15 educational researchers who achieved the degree of MEd or PhD. Two physics teachers whose schools were located in Seoul and their ten students were investigated twice and answered paper survey. Data from the five sources described above was analyzed and categorized according to the three target groups of WEBSYSTEM. Table 2 shows the three target groups (Educational Researcher, Teacher, and Student) in the first column and their related items of the 'Requirement Analysis' in the second column marked with R1-R9. Each of the 'Requirement Analysis' items are linked to several detailed items in the third column additionally marked with "a-d" along with the second column marks, holding their various sources in the fourth column.

The analysis reveals that educational researchers call for a wide range of improvement of this system for testing or investigating physics misconceptions. Furthermore, they suggested that the WEBSYSTEM was an efficacious research tool for developing instructional models (R1-R3). The teachers required instructional systems before (R4), during (R5) and

Table 2
'Requirement Analysis' for WEBSYSTEM

Target group	Requirement	Detail	Source
Educational Researcher	R1. Wide range of study	R1a. International study	Song et al (2004)
		R1b. National large-scale study	Song et al (2004)
		R1c. Study on curriculum	Seminars
		R1d. Longitudinal study on individual/group	Researchers
	R2. Improvement in system for testing or investigating misconceptions	R2a. Study to improve test system	Seminars
		R2b. Study to refine previous tests	Seminars
		R2c. Study to expand the domain of conception test (through multimedia, etc)	Seminars
	R3. Developing instructional models	R3a. Developing instructional models	Song et al (2004)
		R3b. Searching the implementation of the test results	Song et al (2004)
Teacher	R4. Diagnostic evaluation	R4a. Automated diagnosis evaluation	Teachers
		R4b. Investigation of students' conceptions with national wide data	Teachers
		R4c. Investigation of properties on their classes	Teachers
		R4d. Distance education	Teachers
	R5. Instructional data for concept learning	R5a. Instructional data for lesson planning	Taber (2003)
		R5b. Study on concept instruction	Seminar
	R6. Formative evaluation	R6a. Investigation students' conceptual change before and after instruction	Seminar
Student	R7. Distance learning	R7a. Learning out of school	Student
	R8. Self-directed learning	R8a. Preparing school learning	Taber (2003)
		R8b. Reviewing school learning	Teacher
	R9. Recording students' phase of learning	R9a. Recording students' conceptual change	Seminar

after (R6) their physics classes from WEBSYSTEM. The students required and expected WEBSYSTEM to provide opportunities for their self-directed learning in distance education (R7, R8), as well as recording their progress of learning (R9).

2. System Design

The three main functions for WEBSYSTEM were designed according to the items of the 'Requirement Analysis'. 13 chosen items of 'System Design' are shown in Table 3. The three main functions ('Automated test', 'Real-time analysis of the test', 'Feedback with results of the test') and the 13 detailed functions (F1-F13) are presented, which were designed in accordance with the resulting items of the 'Requirement Analysis'.

In order to benefit students and teachers as well as researchers by automation of WEBSYSTEM, the 13 functions of WEBSYSTEM were arranged as shown in Fig. 1. The highly-automated system manipulated by one administrator enabled teachers to instruct students to perform the diagnostic test by themselves at schools or even at their home with great efficiency.

3. System Development

When it comes to Internet-based tools in the field of education, WebCT (online proprietary virtual learning environment system, www.webct.com) is referred as the first commercially successive system, which was introduced first in 1996[1]. Other than WebCT, Sakai Project (online Collaboration and

Table 3

'System Design' for WEBSYSTEM (ER: Educational Researcher, TE: Teacher, ST: Student)

Main function	Detail	Requirement		
		ER	TE	ST
Automated test	F1. Participants' basic information (name, school year, region etc)	R1		R9
	F2. Participants' advanced information (test date, test count, etc)	R1, 6		
	F3. Identifying repeated participation	R1, 6		R9
	F4. Randomizing the sequence of the test	R2		
	F5. Recording response time (seconds)	R2		
	F6. Valid email checkup algorithm	R1		
	F7. Tools to input the data of the test	R3		
Real-time analysis of the test	F8. Test Report 1: Distribution of correct responses and the number of participations according to classroom and school year		R4	
	F9. Test Report 2: Distribution of all responses and the number of participations according to classroom and school year		R4	
	F10. Test Report 3: Distribution of all responses and tables with descriptive responses	R2	R5	
	F11. Real-time bar/percentage charts and tables for analysis		R4,5, 6	
Feedback to students	F12. Test Feedback 1: real-time / email			R7, 8
	F13. Test Feedback 2: Log of participations			R9

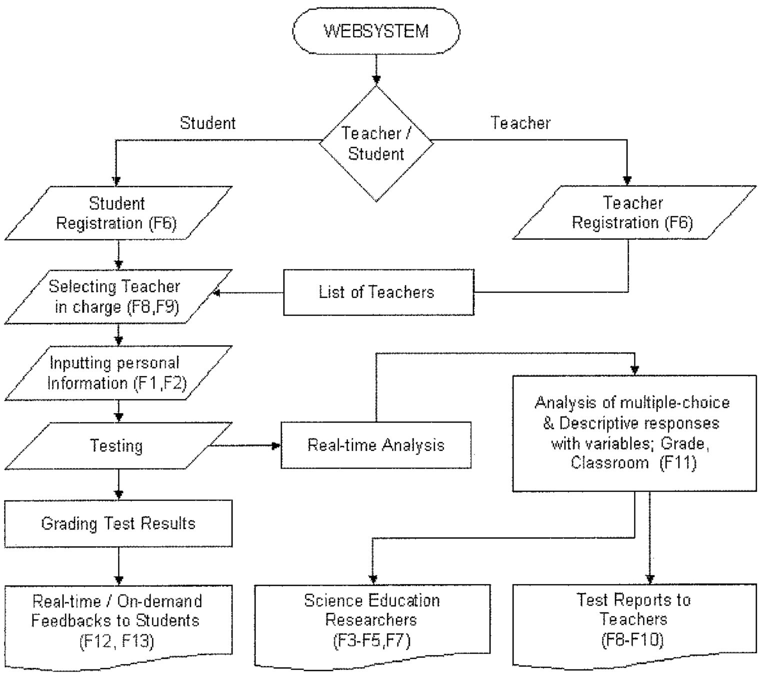


Fig. 1 *Flowchart of functions in WEBSYSTEM*

Learning Environment, www.sakaiproject.org) and Scholar360 (Learner Management System, www.scholar360.com) among others have been adopted for integrated course work. However, this manufactured software failed to fulfill our expectations for the 'Requirement Analysis' and 'System Design', since they are not supportive for modifications of program source by administrators or educational re-

[1] http://en.wikipedia.org/wiki/WebCT

searchers. Even though their functions are modulated to play high flexibility to fit into general educational purpose, they lack many of the requirements (R1 -R3) of researchers of students' misconceptions.

Based on our research findings, the programmer, who was one of the authors, reviewed other web -based programs and adopted three of them as following: 'no1' from the web-based test-bank program (www.sinabro21.com), 'ChartDirector' in the web-based chart program from Advanced Software Engineering (TM) (www.advsofteng.com) and 'zeroboard' from the web-based community program (www.nzeo.com). 'no1' is a type of online evaluation system which is supportive for supplementary modifications contrary to the above-mentioned software. Moreover, 'no1' has non-encrypted data base structure, which means that it is open to cognitive researchers for analyzing and processing students' test results according to their research purpose. These three programs were applied to WEBSYSTEM, and additional functions were coded by one of the authors. The total development period was eight months. The final result of 'System Development' was an Internet site (epic.snu.ac.kr/test) embedding our WEBSYSTEM (See Fig. 2).

4. System Implementation

The authors examined the application of WEBSYSTEM through a case study of teacher K and his 247 students, based on items from the 'Requirement Analysis' toward teachers and students (R4-R9) as well as the results of the 'Wave and Light' misconception test (R1-R3). After reviewing students' responses and an administrative meeting with teacher K, the students participated in the test via the Internet at home. In addition, data was collected through teacher interview, teacher's after-class report, and student paper survey, which were also part of the next procedure of 'System Evaluation'.

1) Administrative meeting with teacher K

At the administrative meeting with teacher K, the authors found out that, in addition to the availability of computer/Internet at school, most students had a computer and easy access to the Internet at home. Teacher K said most students were capable of using the Internet for assignments since they had done so previously in his class. Teacher K, who had conducted diagnosis evaluation with OMR (optical mark reader) cards suggested that it would be time-efficient to take test via the Internet (R4). Thus, the authors could confirm the presence of critical factors such as students' capability in using the computer/Internet, scheduling, and administrative strategies to increase students' participation.

2) Real-time analysis of students' test results

Out of 247 students who had been informed via WEBSYSTEM through which the test was given as homework by teacher K, 161 10^{th} graders (48% of the total 10^{th}s) and 86 11^{th} graders (64% of the total 11^{th}s) participated in the test via WEBSYSTEM over a period of 12 days. Teacher K and the authors examined the test results by three real-time 'Test

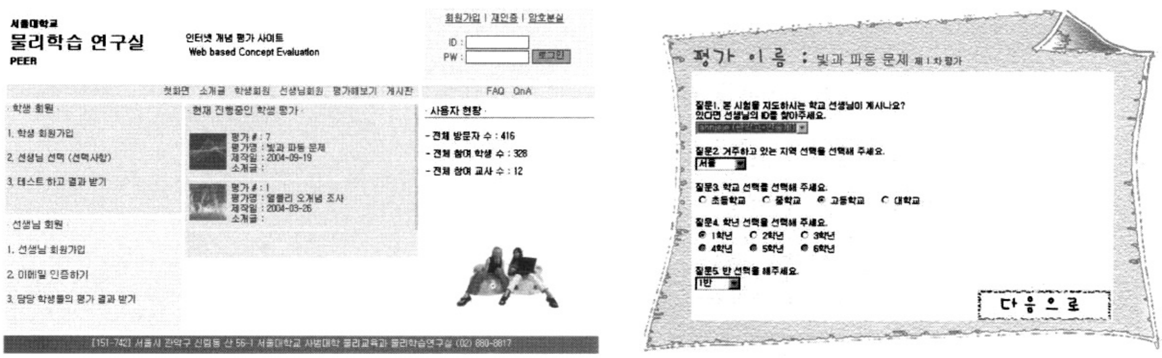


Fig. 2 *Screenshot of Main page (Left) and Testing page (Right). The main page presents instructions for registrations, a brief summary of on-going tests, and the number of current participants. After visiting the main page, participants encounter the testing page where students input their personal information.*

Reports'. WEBSYSTEM automatically provided five real-time analyses of the test:

a) Test Report 1 contained test information, distribution of correct responses and the number of participations in classroom code, school year, and links to the other 'Test Reports 2, 3' (See Fig. 3)
b) Test Report 2 contained questions, scientific answers, explanations, and the distribution of all responses, the number of participations in classroom code, and school year (See Fig. 4)
c) Test Report 3 contained questions, scientific answers, explanations, the distribution of all responses, and a table with descriptive responses in selected classroom code, and school year (See Fig. 5)
d) Test Feedback 1 contained questions, scientific answers, explanations, students' responses delivered through email. Since WEBSYSTEM confirms students' email address as one procedure of registration, students could be informed their test results through two types of services, real-time or on-demand emailing service even without visiting the system. Teachers decide when students access their test results in order to employ 'Test Feedback 1' as a teaching and learning material in the science class.
e) Test Feedback 2 contained history of each participation (i.e., test information, date/time) classified by students' plural participations. Participants can review their phase of learning by recognizing their conceptual change with a time variable by visiting WEBSYSTEM.

Since teachers consider differences among heterogeneous students grouped by classrooms, they use it to determine instructional strategies for their science classes. WEBSYSTEM was designed to provide real-time data analysis, 'Test Report 1', comparing the mean score per classroom and school year. As an example, Fig. 3 indicates that 161 first graders in middle school (= grade 10 students) participated in the 'Wave and Light' misconception test and the average score (MS) for the 10 different classrooms is 34. First graders in classroom code 'Class 4' were 25 participants (n=25) and had the lowest average scores (MS=28). The responses for each question from '1-4' could be compared to ones of other classes in detail through 'Test Report 2' (See Fig. 4).

No.	Test Information	Status of participants
	Sequence: 7 Title: Wave and Light Initiation date: 19-9-2004 Profile:	* Year 1 (n=161, MS=34) -Class 1 (n=20, MS=29) -Class 2 (n=10, MS=33) -Class 3 (n=20, MS=32) -Class 4 (n=25, MS=28) -Class 5 (n=11, MS=38) -Class 6 (n=17, MS=39) -Class 7 (n=10, MS=39) -Class 8 (n=21, MS=33) -Class 9 (n=9, MS=37) -Class 10 (n=18, MS=39) * Year 2 (n=86, MS=41) -Class 7 (n=14, MS=41) -Class 8 (n=22, MS=38) -Class 9 (n=21, MS=42) -Class 10 (n=29, MS=44)

Fig. 3 *Translated screenshot of 'Test Report 1', Columns indicate test number, test information (sequence, title, date of initiation, and profile), and state of the participations (per school year, and classroom code). For better understanding, the original screenshot written in Korean was translated in English.*

Fig. 4 shows the result from the question number 1 ('A moving wave suddenly changed its speed. What made this happen?', i1=Change in its frequency, i2=Change in its width, i3=Change in thickness of the rope, i4=Change in its height) with Year 1 participants. The scientific concept is known as i3. The x-axis arranges ten classroom codes from 1 to 10 where each of them has a number of participants in brackets; for example, 25 students are from a group of classroom code 4. In this specific group, 37.50% students have responded on the item, 'i1' of misconception. Detailed information for 'student name', 'response', 'descriptive response', and 'time of response' is shown in 'Test Report 3' (See Fig. 5).

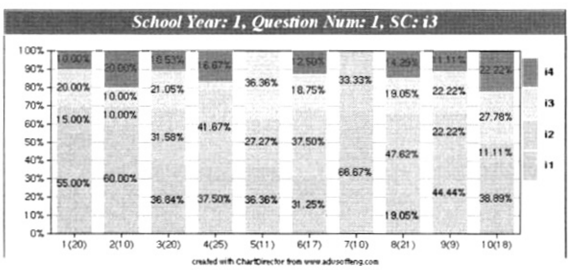
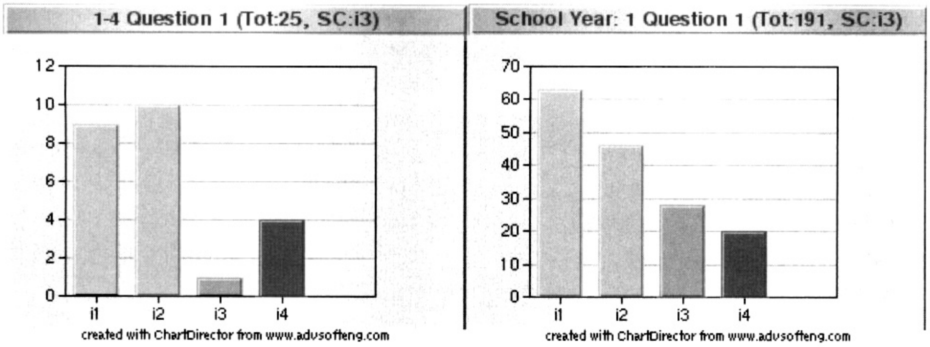


Fig. 4 *Screenshot of 'Test Report 2' indicating distribution of multiple-choice answers from first graders grouped by ten classes on question number 1 with scientific concept i3*



Name	Answer	Descriptive Answer	Response Time (s)
Student A	i2		192
Student B			127
Student C	i4	When wave fells down from the height it seems to get faster.	1336
Student D	i1	I saw it in WinAmp (mp3 playing software).	1260
Student E	i4	I don't know.	49
Student F	i2		31
Student G	i1	It gets faster because of the frequency.	74
Student H	i1	The stronger frequency made it faster.	64
Student I	i4	The height of the graph seems to be an indicator of the velocity.	28
Student J	i1	Change in the frequency causes change in wave.	141

Fig. 5 *Screenshot of 'Test Report 3' indicating distribution of multiple-choice answers and a table of descriptive answers from the first graders of classroom code '4' on question number 1. The lower table from the original screenshot written in Korean was translated in English*

While 'Test Feedbacks 1, 2' aim to present teachers overall distribution of the mean score per classroom and school year, 'Test Feedback 3' provides detailed information of each students including their multiple-choice answers, descriptive answers, and response times in second, as shown in Fig. 5. In this specific example of 25 first graders in classroom code '4' with question number 1, less proportion of students choose i3 of the scientific concept in comparison to one of total 192 first graders. By analyzing students' descriptive answers asking the reason of their multiple-choice answers, teachers can browse individual students' misconceptions. For example, Student C in Fig. 5 attributed the increase in velocity of a moving wave to the height that enacts the wave to be attracted by the gravity. Teachers can always access WEBSYSTEM and be provided with such specific information of students' conceptions to employ in their science class.

'Test Feedbacks 1,2' consist of identical components of the 'Test Report 3', additionally including scientific explanations of each item that were adopted from Song *et al.* (2004)'s publication of a physics misconception test set for this prototype research. For participants' conceptual studying material, WEBSYSTEM publishes 'Test Feedbacks 1, 2' through real-time or on-demand emailing service. Students are informed of their test grades, scientific explanations of each item, their responses and the record of date/time.

5. System Evaluation

The last procedure in ADDIE is to evaluate by summarizing and documenting feedbacks from the teachers and students regarding the effectiveness of the system, so that improvements can be made the next time of implementation (Gustafson, Branch, 2002). Analysis of feedbacks including teacher interviews, a teacher's after-class report and a student survey narrates the feedbacks from the practical implementations and summarizes every earlier step, particularly concerning whether or not the requirements and functions have been accomplished through the system.

1) Teacher interview

According to the interview questionnaire (shown as Table 1), Teacher K was interviewed concerning three respects: "experience in class", "analysis of test results" and "opinions of system". Firstly for his experience in class, through teacher K's prior experience with students' internet capabilities, the authors found that his students had been accustomed to this kind of testing on internet, since none of 247 participating students asked how to access or participate in it. He showed us a document that he had made with a specific guide of the WEBSYSTEM, and said that students' care of the present grading system emphasizing students' eagerness in science class might have increased their participation.

Secondly, through his analysis of the test results, the Teacher K became aware of the differences in each of the 14 classes and could prepare different instructional content for each of them (R4). He would adopt the students' descriptive responses as reference when nominating student discussion leaders who were shown distinctive misconceptions in 11^{th} graders' physics classes (R5). Because students' descriptive responses are written in the students' own words and are easy to recognize, another possible usage of the students' descriptive responses suggested by teacher K is to derive multiple-choice questionnaire items for their final examination with them.

Lastly, for his opinions of system, Teacher K suggested that the teacher's opinion should be made into reviewing and selecting of test items, which we regard as another administrative strategy of WEBSYSTEM.

2) Teacher's after-class report

After Teacher K's class on 'Wave and Light', authors asked him to submit his after-class report regarding three issues. On the issue of 'Comments on usage of WEBSYSTEM in physics class', he indicated that it had worked as a useful tool for investigating students' misconceptions before class, and it provided data for comparison between classes (R4, R5).

> It is very useful to have information of students' misconception before their science class. WEBSYSTEM provides me concrete data of their cognitive knowledge about what they would learn among classes, which I have been guessing beforehand. (Teacher K)

Regarding the second issue, 'Examples of usage on WEBSYSTEM in physics class', He exampled the question number 4 ('How does a plane wave on a barrier containing an opening spread afterward?') where WEBSYSTEM provided him with efficacious information to predict students' behavior. He mentioned below, which is proved to be relevant with (R4):

> Students in high-scoring classroom show little interest in their science class, while students in low-scored classroom seemed more motivated and were more active. (Teacher K)

On the final issue, 'Difficulties of usage on WEBSYSTEM in physics class', he suggested that the question items should be reexamined for their relevance to school science curriculum, because some students misunderstood the situation of some questions. Although we had employed Song *et al.* (2004)'s standardized misconception test, the issue of validating the test items had been raised through our case study.

> The test material needs to be reviewed properly concerning its validity for secondary students. For example, reflected light from surface of mirror decreases its intensity. Strictly speaking, however, the question in the system remarks both densities of incident and reflected light are identical. That could be wrong. (Teacher K)

Turning to another discussion, Teacher K asked us for data about the percentage of all responses per question per year that had already been included of 'Test Report 3'. This misunderstanding indicated a need for more thorough teacher orientation about the user interface of WEBSYSTEM.

3) Student survey

Ten students were asked about 'Technical difficulties', 'Benefits of WEBSYSTEM', and 'Risks of WEBSYSTEM' using the paper survey. Results indicated that students had no technical difficulties with WEBSYSTEM. One student wrote, 'People wouldn't find any problem on this web site these days'. Secondly, five students said that students might benefit from WEBSYSTEM, by 'Checking and confirming what he/she has to know before the class. On the other hand, five students expressed that not having 'Test Feedbacks' by real-time email right after their testing was inconvenient, which opposes teacher K's request to use on-demand email. WEBSYSTEM has been equipped with such settings of real-time / on-demand emailing feedback. This finding arouses the importance of administrative strategies of when to provide students their 'Test Feedbacks'.

4) Summative system evaluation

The evaluation sources (teacher interviews, teacher's after-class report, and student survey) were summarized into the final procedure of evaluation, "summative system evaluation". Since the complex web-based system has functions of infrastructures (F1-F7) as well as functions of outer interfaces (F8-F13) that teachers or students can access only in limited scopes, the summative system evaluation was conducted as a type of self-reporting, synthesizing the feedbacks during the whole procedures through the system evaluation procedure. Holding the "Requirement Analysis" as criteria, this summative procedure was performed by the author of a programmer who technically understands every structure of the functions and infrastructures. The findings are shown in Table 4.

Most of the requirements from the Educational Researcher (R1a-R1c, R2a, R2b, R3a) were succe-

Table 4
'System Evaluation' for WEBSYSTEM (A: Successful, B: Partially Successful, C: To be revised, D: To be developed)

Target group	Requirement	Detail	Grade
Educational Researcher	R1. Wide range of study	R1a. International study	B
		R1b. National large-scale study	A
		R1c. Study on curriculum	B
		R1d. Longitudinal study on individual/group	C
	R2. Improvement in system for testing or investigating misconceptions	R2a. Study to improve test system	A
		R2b. Study to refine previous tests	A
		R2c. Study to expand the domain of conception test (through multimedia, etc)	D
	R3. Developing instructional models	R3a. Developing instructional models	B
		R3b. Searching the implementation of the test results	B
Teacher	R4. Diagnostic evaluation	R4a. Automated diagnosis evaluation	A
		R4b. Investigation of students' conceptions with national wide data	A
		R4c. Investigation of properties on their classes	A
		R4d. Distance education	A
	R5. Instructional data for concept learning	R5a. Instructional data for lesson planning	A
		R5b. Study on concept instruction	A
	R6. Formative evaluation	R6a. Investigation students' conceptual change before and after instruction	C
Student	R7. Distance learning	R7a. Learning out of school	A
	R8. Self-directed learning	R8a. Preparing school learning	B
		R8b. Reviewing school learning	B
	R9. Recording students' phase of learning	R9a. Recording students' conceptual change	A

ssful or partially successful with grade A or grade B. The reason for this success was based on the fact that their relevant functions (F1, F2) succeeded in operating without any error to generate a database with basic/advanced information containing over 30 tables in the online database system. Whereas, the requirement for the 'Longitudinal study' (R1d) received a grade of grade C since functions (F2, F3, F11) to output test results with a time variable needed to be revised. The intension of extending the domain with multimedia applications (R1d) also failed to generate any linked functions, which thus led to grade D.

On one hand, teachers' requirements of diagnostic (R4a-R5b) were all evaluated with grade A on the grounds of the investigation (teacher interviews and student survey). On the other hand, for the reasons above that the requirement of longitudinal study (R1d) lacks of time variables functions, requirements for 'Formative evaluation' (R6a) is evaluated to need to be revised, resulting in grade C.

Functions of the automated test and feedbacks (F1-F7, F12-F13) were described to be designed adequately, which thus leads to a grade A on 'requirements for distance learning and their recording' (R7, R9). According to evidence from the qualitative survey presenting 'no technical difficulties with system implementation', the confliction between the teacher's request for blind tests not to distribute the detailed test information among students in advance and students' request for real-time email with their test feedbacks resulted in grading the requirement of self-directed learning (R8) as a partial success, grade B. This issue needs to be dealt in scope of the administrative strategy, which seems to be one of the critical issues in the implementation of WEBSYSTEM.

V. Research Limitation

The summative evaluation basically refers to our qualitative research findings based on interviews and an after-class report, which were reviewed according to the programmer's insight of the requirements and functions in WEBSYSTEM. As the author of a web-based programmer performed synthesizing evaluative data into summative evaluation, an issue of validity in the summative evaluation of system (shown in Table 4) needs to be discussed. This evaluating procedure, which depends on the proficiency of the programmer, was adopted to prevent an accomplished requirement from being misevaluated, when teachers or students misunderstand some functions as seemingly mal-functioning or non-existing. While conventional definitions of validity for evaluations rely only on the technical side to understand assessment, literature suggests that we also have to understand what happens when assessment results are used (Dylan, 2003). In the light of such specific concept of validity, the summative evaluation in our study can be authorized for certain usages of 1) providing practical information for cognitive researchers who adopt WEBSYSTEM and 2) enhancing the system by employing requirements graded as C or D into requirements of a following system version for performing developmental cycles.

VI. Conclusion & Discussion

We have developed a web-based system for testing students' misconceptions in physics employing the instructional building model, ADDIE, and the requirements-driven methodology. With the detailed description, the five procedures of ADDIE have clarified the development of WEBSYSTEM to provide educators and science educational researchers with abundant information for reproducing or improving the web-based system.

The implementation of WEBSYSTEM into a science teacher and 247 students revealed that it could perform effectively as a multi-purposed tool for the three target groups through revisions on time-variable functions (See Table 2 & 4.). With secondary analysis, the data from students' test results stored in the database could provide opportunities for a wide-ranged and detailed investigation on physics misconception and its applications (R1-R3). Besides the requirements derived from teachers (R4-R6), another possible usage of WEBSYSTEM revealed by the teacher interview was to process students' descriptive conception into multiple-choice questionnaire items. Such multiple-choice items will capture students' presentation on science conception with statistics analysis better than paragraph responses, and with better clarity than Likert-type responses (Aikenhead, Ryan, 1992). WEBSYSTEM also turned out to operate adequately enough to assist students in benefiting from distance learning (R7), self-directed learning (R8), reviewing their learning phase (R9), as found in the procedure of system evaluation. For students, they were provided with formative tests through WEBSYSTEM at their home or school before their science classes. The automated feedbacks from WEBSYSTEM via email would enable students to perform better in their schoolwork. Since WEBSYSTEM provides direct authentic evaluation experience for students, the automated feedbacks via email would encourage their capacity to monitor the quality of their own work during actual science classes (Sadler, 1989).

In our study, one of the most significant difficulties is insufficient proficiency of programming the Internet-based system, since one of the authors, who does not possess a degree or a career in educational engineering, conducted the entire process of the system development by himself as a novice programmer. Consequently, three requirements, "R1d. Longitudinal study on individual/group", "R2c. Study to expand the domain of conception test through multimedia, etc.", and "R6a. Investigation students' conceptual change before and after instruction" were evaluated to-be-revised or to-be-developed. An issue of the improper user-interface issue was also interviewed from Teacher K. Moreover, WEBSYSEM is not as esthetical or full of up-to-date technologies as other predominant educational web sites where professional designers, programmers and server administrators cooperate systematically. These shortcomings were mainly due to a lack of such proficiency.

With all these defects, the prototype research through the rationales from the educational engineering has provided practical information of designing, developing and practicing an Internet-based research system in detailed for educational researcher or schoolteachers. Our study also presented a research paradigm in which three heterogeneous targets (science educational research, schoolteacher and students) evenly benefit from the web-based System for Testing Students' Physics Misconceptions (WEBSYSTEM). Our suggestions listed below could improve the system, refining its requirements and functions to lead better benefits for the three target groups.

a) For better implementation of WEBSYSTEM into schools, follow-up studies should be made concerning teacher orientation.
b) Attempts to examine the misconception tests and extension of its methodological domain should be conducted for standardization of misconception testing.
c) International investigations of students' science conceptions involving students' cultural context need to be conducted via WEBSYSTEM.

Acknowledgement

The authors express their special thanks to Jan Sølberg from The Danish University of Education, Denmark and Jari Lavonen from University of Helsinki, Finland. These science educational researchers from Europe were eager to review and discuss the article during their stay with the author in Kobe, Japan.

References

Aikenhead, G. S., & Ryan, A. G. (1992). The Development of a New Instrument: Views on Science-Technology-Society (VOSTS). Science Education, 76(5), 477-491.

Andriole, S. J. (1997). Requirements-Driven ALN Course Design, Development, Delivery & Evaluation. Journal of Asynchronous Learning Networks, 1(2), 57-67.

Chi, M. T. H., Slotta, J. D., & Leeuw, N. D. (1994). From things to processes: A theory of conceptual change for learning science concepts. Learning and Instruction, 4(1), 27-43.

Ministry of Education & Human Resources Development (2003). Survey of secondary school science teaching-learning materials: Ministry of Education & Human Resources Development.

Dylan, W. (2003). Validity: all you need in assessment. School Science Review, 85(311), 79-82.

diSessa, A. A. (2002). Why conceptual ecology is good idea. In M. Limon & L. Mason (Eds.), Reconsidering conceptual change: Issues in theory and practice (pp. 29-60). Dordrecht: Kluwer.

Engelbrecht, J., & Harding, A. (2004). Combing Online and Paper Assessment in a web-based Course In Undergraduate Mathematics. Journal of Computers in Mathematics and Science Teaching, 23(3), 217 -231.

Fensham, P. J. (2004). Defining an Identity: The Evolution of Science Education as a Field of Research (Vol. 20). Dordrecht: Kluwer Academic Publishers.

Korea Science Foundation (2003). Study of the National Survey on the Students' Recognition toward Science & Technology: Korea Science Foundation.

Gay, L. R. (1996). Educational research: Competencies for analysis and application (5th ed.). Comumbus OH: Merrill Publishing Compay.

Gustafson, K. L., & Branch, R. M. (2002). Survey of Instructional Development Models (4th ed.). New York: Eric Clearinghouse on Informaion & Techonlogy.

Kim, I. G. (1991). College Students' Conceptual Change about Force and Acceleration through Critical Discussion of the Rival Concepts based on Evidences and Reflective Thinking. Seoul National University, Seoul.

Mertler, C. A. (2003). Patterns of response and nonresponse from teachers to traditional and web surveys. Practical Assessment, Research & Evaluation, 8(22).

Oh, W. K. (1998). Secondary students' conceptual change in force and motion according to the structured contrastive activity by introducing weightlessness and relativity as its non-everyday contexts. Seoul National University, Seoul.

Oshima, J., Oshima, R., Murayama, I., Inagaki, S., Takenaka, M., Nakayama, H., et al. (2004). Design experiments in Japanese elementary science education with computer support for collaborative learning: hypothesis testing and collaborative construction. International Journal of Science Education, 26(10), 1199-1221.

Park, J., Seo, J., Chung, B., & Pak, S. (1994). An Analysis of Middle School Student's Responses to the Deductive` Reasoning Task for Change of Concept about Force and Motion. Korea Association for Research in Science Education, 14(2), 133-142.

Park, J.-W. (1996). Internet Use in Network-based Science Education. The Journal of the Institute of Science Education (Chongju National University of Education), 17.

Park, S.-T., Lee, H., & Yuk, K.-C. (2003). Exploring Applications of 3D web-based Virtual Reality Technology in Physics Education. Sae Mulli, 46(4), 179-186.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Towards a theory of conceptual change. Science Education, 66(2), 211-227.

Rabinowitz, S., & Brandt, T. (2001). Computer -based Assessment: Can It Deliver on Its Promise? Knowledge Brief. San Francisco: WestEd.

Rha, I., & Chung, H. (2001). Developing an Action Model for WBI Design. The Korean Society for Educational Technology, 17(2), 27-52.

Russell, M., & Tao, W. (2004). The influence of computer-print on rater scores. Practical Assessment. Research & Evaluation, 9(10).

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science 18(2).

Savenye, W. C., & Robinson, R. S. (2003). Qualitative research issues and methods: An introduction for educational technologists. In D. H. Jonassen (Ed.), Handbook of research on educational communications and technology (2nd ed.). New York: Macmillan Library Reference USA.

Shim, K.-C., Kim, H.-S., & Chung, J.-I. (2005). The Effect of web-based Learning by Studying the Motion of the Moon. Korea Association for Research in Science Education, 25(4), 450-464.

Song, J., Kim, I. G., Kim, Y. M., Kwon, S., & Oh, W. K. (2004). Map of Students' physics Misconceptions. Seoul: Book's-hill.

Taber, K. S. (2003). Responding to alternative conceptions in the classroom. School Science Review, 84(308), 99-108.

Tao, P.-K., & Gunstone, R. F. (1999). Conceptual change in science through collaborative learning at the computer. International Journal of Science Education, 21(1), 39-57.

Vosniadou, S. (2002). On the nature of naive physics. In M. Limon & L. Mason (Eds.), Reconsidering conceptual change: Issues in theory and practice (pp. 61-76). Dordrecht: Kluwer.

Wade, V. P., & Lyng, M. (2000). An automated Evaluation Service for Educational Courseware. Paper presented at the World Conference on the

WWW and Internet, San Antonio, Texas, USA.

Zele, E. V., & Lenaerts, J. (2004). Improving the usefulness of concept maps as a research tool for science education. International Journal of Science Education, 26(9), 1043-1064.

Zietsman, A. I., & Hewson, P. W. (1986). Effect of instruction using microcomputer simulations and conceptual change strategies on science learning. Journal of Research in Science Teaching, 23(1), 27-39.