What strategies are effective for formative assessment in an e-learning environment?

T.H. Wang

Department of Education, National Hsinchu University of Education, Hsinchu City, Taiwan

Abstract

The web-based formative assessment developed in this research is named Formative Assessment Module of the Web-based Assessment and Test Analysis System (FAM-WATA). FAM-WATA is a multiple-choice web-based formative assessment module containing six effective strategies: 'repeat the test', 'correct answers are not given', 'query scores', 'ask questions', 'monitor answering history', and 'all pass and then reward'. This research explored the effectiveness of FAM-WATA, cognitive styles and e-learning, and student attitudes towards the six strategies of FAM-WATA. A total of 503 seventh-grade students in central Taiwan were valid in this research. Overall results indicated that students displayed a positive attitude towards the six strategies of FAM-WATA. In addition, results also showed that students in an e-learning environment equipped with FAM-WATA achieved better learning effectiveness, and that field independent students appeared to make better use of FAM-WATA strategies than field dependent students. This research concluded that FAM-WATA benefited student learning in an e-learning environment.

Keywords

cognitive style, e-learning, FAM-WATA, WATA system, web-based formative assessment.

Introduction

With the advent of the World Wide Web and information technology (IT), it has become easier to implement an e-learning curriculum. An e-learning environment facilitates student learning without the constraints of time and distance, giving students more opportunities to control their own learning. However, Sujo de Montes and Gonzales (2000) stated that online students tended to be more risk takers in terms of technology use and technology integration, and suggested that learners in the e-learning environment need to be their own teachers. However, self-teaching in an e-learning environment often makes learners feel isolated and dis-

connected because of the lower level of teacher supervision in the e-learning environment. This is especially true in Taiwan, where education is traditionally teacher-centred and students are poor at taking the initiative in their own learning. Thus, stimulating student motivation and making them learn more actively has become an important issue in the field of e-learning.

Bransford *et al.* (2000, pp. 139–144) have observed that an assessment-centred teaching environment is an effective design. They further pointed out that formative assessment plays an important role in an assessment-centred learning environment and motivates students to learn and directs their learning. Although formative assessment is very important in a learning environment, current knowledge about embedding formative assessment in learning environments is lacking (Bell & Cowie 2001). Formative assessment is likely to be one of the most important research fields in the future (Bransford *et al.* 2000; Bell & Cowie 2001, p. 257).

Building on previous research (Wang *et al.* 2004a), this research develops and explores web-based formative assessment strategies in an e-learning environment. Using the Web-based Assessment and Test Analysis (WATA) system (Wang *et al.* 2004a) as a foundation, the Formative Assessment Module of the WATA system (FAM-WATA) was developed and implemented in an e-learning environment to help teachers interact with students and give feedback. This system also allows students actively to assess themselves. Furthermore, this research explored the effectiveness of FAM-WATA in junior high school e-learning environment. Previous research has shown that cognitive style is known to be an important determinant of student learning in an e-learning environment (Kim 2001; Chen *et al.* 2005). This research also used the cognitive style approach to investigate the effectiveness of three different types of formative assessment in an e-learning environment.

Literature review

Formative assessment

Ebel and Frisbie (1991) observed that the terms 'formative' and 'summative' were introduced by Scriven (1967) to describe the various roles of evaluation in curriculum development and instruction. Ebel and Frisbie defined 'summative assessment' as assessment conducted at the end of instruction to determine if learning is complete enough to warrant moving the learners to the next segment of instruction. Thus, summative assessment is always used to investigate student learning achievement, and is always administered at the end of the teaching process. Ebel and Frisbie stated that 'formative assessment' is conducted to monitor the instructional process and to determine whether learning is taking place as planned. However, the purpose of the formative assessment during the teaching process is to illuminate learner difficulties and enhance teacher effectiveness, and is always administered during the teaching process.

Another important function of formative assessment is providing students with 'continuous feedback', meaning that opportunities for feedback should occur continuously, but not intrusively, as a part of instruction (Bransford *et al.* 2000, p. 140). Bell and Cowie (2001) concluded that formative assessment is increasingly being used to refer only to assessment that provides feedback to students (and teachers) about learning occurring during the period of instruction and learning, and not after. Feedback from formative assessment is beneficial in the adjustment of teaching strategies and application of appropriate remedial techniques. Bell and Cowie suggested that the teacher gather assessment information (feedback) about student learning from formative assessment, and then respond to promote further learning, and then shape and improve student competence (Gipps 1994).

Feedback in formative assessment can uncover weaknesses requiring reinforcement and is seen as an essential component of the formative assessment interaction, where the intention is to support learning (Sadler 1989; Perrenoud 1998). Teachers should make use of formative assessment to give students feedback at an appropriate point in the learning process (Brown & Knight 1994) because formative assessment will be of little help to learners if teachers do not allot time for feedback. Wiliam and Black (1996) observed that messages from the feedback of formative assessment should tell learners what must be done, thus improving learning effectiveness. In traditional classroom teaching, teacher feedback to learners is limited (Bransford *et al.* 2000, pp. 140–141). Typical teachers give just one summative assessment, like transcripts or assignment scores. After grades are given, students move on to the next topic and work for another grade. In other words, learning is driven by reading and making grades. Such learning is not effective. Instead, the truly meaningful feedback that teachers give to learners should derive from formative assessment. By using meaningful feedback, students can improve weaknesses in learning and thinking, increase and transfer learning, and value opportunities to revise (Barron *et al.* 1998; Black & Wiliam 1998; Bransford *et al.* 2000, p. 141).

However, formative assessment feedback is difficult for teachers to provide, because they face large numbers of students, lengthy pieces of work, or practical constraints such as time and workload (Buchanan 2000). In this research, a web-based formative assessment system, FAM-WATA, was developed and used to address this problem. This research not only applied FAM-WATA to assist teachers in giving feedback and interacting with students in an e-learning environment but also explored the effectiveness of FAM-WATA in facilitating student e-learning effectiveness.

Web-based formative assessment strategy design

Bell and Cowie (2001), Bransford *et al*. (2000, pp. 140–141), Perrenoud (1998), Sadler (1989), and Williams (2004) observed that feedback is as an essential component of successful formative assessment. Brown and Knight (1994), Buchanan (2000) and Wiliam and Black (1996) also pointed out that the feedback should be 'timely feedback' which meets the requirements for formative use. Bransford *et al*. further suggested that teachers should make use of formative assessment to provide students with opportunities to gain feedback and revise their thinking. Thus, it is important for teachers to provide learners with opportunities for receiving 'timely feedback' and 'repeating the test'.

Henly (2003) and Khan *et al*. (2001) embedded the design of 'repeat the test' and 'timely feedback' in their web-based formative assessment and found significant effects on learning effectiveness. Learners discovered what they had missed and were motivated to clarify the concepts to master content. Buchanan (1998, 2000) found that while 'repeat the test' was an important strategy in an e-learning environment, the functions of 'correct answers are not given' and 'timely feedback (swift and useful feedback)' will make web-based formative assessment more effective. In other words, when students have the opportunity to repeat the test, correct answers should not be given. Moreover, the system should give references to the answers when students answer incorrectly. Buchanan (2000) showed that these strategies enabled learners to engage more fully with the course materials. He advised repeating the test after doing the suggested reading in a test–learn–retest cycle, which continues until the subject matter is mastered. These strategies will help participants know which items need reinforcement, and enable learners to concentrate on learning materials and steadily master course concepts.

Web-based formative assessment and e-learning effectiveness

Many studies (see Table 1) have obtained positive results with the application of formative assessment to e-learning environments (Buchanan 2000; Khan *et al*. 2001; Gardner *et al*. 2002; Peat & Franklin 2002; Velan *et al*. 2002; Henly 2003; Brewer 2004; Justham & Timmons 2005).

Cognitive style

Cognitive style is usually described as a personality dimension that influences the way individuals collect, analyse, evaluate, and interpret information (Harrison & Rainer 1992; Chen *et al*. 2005). Cognitive style had been studied extensively since the 1970s in an attempt to understand the varying ways that learners perceive and interact with instructional settings, methods, and media (DeTure 2004). There are a variety of dimensions of cognitive styles. DeTure (2004) concluded that the most widely investigated cognitive style is Witkin's 'field dependent/field independent' (Witkin & Asch 1948; Witkin 1950).

Chen and Macredie (2004) concluded that the differences between 'field dependent' and 'field independent' individuals can be categorized into three types: 'global vs. analytical', 'external vs. internal', and 'passive vs. active'. Looking at the 'global vs. analytical' type, Witkin *et al*. (1977) found that field dependent individuals have global perceptions, whereas field independent individuals are good at analytical thought. Chen and Macredie explained that field dependent individuals typically perceive objects as a whole and approach a task more holistically, but field independent individuals focus on individual parts of the object and tend to be more serialistic in their approach to learning. With regard to the 'external vs. internal', Goodenough (1976) found that field dependent individuals rely more on external references; conversely, field independent individuals rely more on internal references. Chen and Macredie revealed that the preference of external or internal references affect individual performance on cognitive restructuring tasks. Field dependent individuals are more strongly influenced by format–structure, whereas field independent individuals are less affected by format–structure (Jonassen & Grabowski 1993). With respect to the 'passive vs. active', Frank and Keane (1993) found that field dependent individuals prefer passive cognitive strategies, but field independent individuals are likely to use active cognitive strategies. Other studies showed that field dependent individuals concentrate relatively first on the whole picture of subjects, rely more on others and external environment, and are less autonomous in cognitive restructuring tasks; conversely, field independent individuals concentrate on one thing at a time, rely on internal cues, and are more autonomous in cognitive

Table 1. Effectiveness of web-based formative assessment in e-learning.

Studies	Subjects	Material	System	Findings
Brewer (2004)	University students	Introductory Biology	BioBytes	Web-based formative assessment helped to improve students' overall understanding of biological principles and concepts.
Buchanan (2000)	Undergraduate students	Introductory Psychology	PsyCAL	Web-based formative assessment can help elevate student learning interests and effectiveness.
Gardner et al. (2002)	University students	N/A	CECIL, Self-assessment	Web-based formative assessment allowed students to enter the environment and practise the questions at any time and from any Internet-linked computer after e-learning. In addition, it also had significant effects on e-learning. Students loved to use the function to learn on the Web by themselves.
Henly (2003)	Dentistry students	Introductory Biochemistry and Molecular Biology	WebCT[1]	Students rated the web-based formative assessment highly, with 80% regarding it as being helpful for their learning.
Justham and Timmons (2005)	Post-registration nursing students	Statistics	WebCT[1]	Web-based formative assessment is a valid and acceptable method of improving knowledge and understanding of statistics within this group of students (post-registration nursing students).
Khan et al. (2001)	Undergraduate medical students	Gynaecology and Obstetrics	Questionmark Perception[2]	Web-based formative assessment allowed individual students to monitor their educational progress and directed their learning.
Peat and Franklin (2002)	University students	Biology	SAM (Self-Assessment Module) developed by Authorware	Students made significant use of the web-based formative assessment; web-based formative assessment helped students in their learning
Velan et al. (2002)	Undergraduate medical students	Pathology	Questionmark Perception[2]	Web-based formative assessment appeared to have been effective in promoting student learning.

[1]http://www.Webct.com
[2]http://www.questionmark.com

restructuring tasks (Chen & Macredie 2004; DeTure 2004).

Some researches have suggested that the development and evaluation of web-based applications should take into account different cognitive styles (Ghinea & Chen 2003; Chen & Macredie 2004; Chen *et al.* 2005). Chen and Macredie observed that web-based instructional programs are used by a population of learners who have different preferences, skills, and needs, which results in new challenges for instructional design. Zoe and DiMartino (2000) argued that further investigation into how diverse populations are using web-based instructional programs is necessary. Hence, in addition to developing a web-based formative assessment strategy, this research also explored its effectiveness in both field dependent and field independent individuals.

FAM-WATA

The WATA system, developed by Wang *et al.* (2004a), is currently equipped with two major modules, SAM-WATA (Summative Assessment Module of the WATA system) and FAM-WATA. SAM-WATA helps teachers administer multiple-choice summative assessment, item analysis, and test analysis, so that teachers can improve teaching, understand the quality of test items, and obtain a snapshot of student learning progress.

FAM-WATA, on the other hand, helps teachers make multiple-choice formative assessment on the Web and construct an assessment-centred e-learning environment. Learners can use FAM-WATA to challenge and evaluate themselves immediately at any time and from any Internet-linked computer without limitation. FAM-WATA offers six main strategies (Wang *et al.* 2004b):

Strategy 1–3: 'Repeat the test', 'correct answers are not given', and 'ask questions' strategies
The combination of two strategies, 'repeat the test' and 'correct answers are not given', in web-based formative assessment will increase e-learning effectiveness (Buchanan 2000). The major purpose of these strategies is to provide students with opportunities to revise the mistakes they have made. In addition to these two strategies, the FAM-WATA tries to stimulate student interest and desire for new challenges through the design of the Web environment, as explained next.

When learners log in and perform a self-assessment, FAM-WATA will automatically choose some questions randomly from the database. The order of questions and options are randomly arranged. This is to prevent learner boredom with repeated tests. A given test item will not show up on the following test if a learner correctly answers the test item three times consecutively. Thus, the number of test items will gradually decrease with each iteration of the test. At some point, all questions will be answered correctly, and the system will tag the successful learner with a 'pass the test' mark. By the same token, if learners cannot answer the test item correctly three times consecutively, then the answer count will be reset to zero and begun again. Answering a test item correctly three times consecutively is necessary because the system judges that the learners may answer the question correctly simply by guessing. The purpose of this design is for learners to actively take on the challenge of learning, not passively guess their way through.

In the above design, 'timely feedback' is combined to form the strategy of 'correct answers are not given'. After learners submit their test papers, FAM-WATA will immediately give scores and present references to learners without directly giving the correct answers of the questions. Meanwhile, learners may also asynchronously interact with teachers by asking questions online. As for the function of 'timely feedback', the system offers learners reference materials to help them find correct answers.

Strategy 4: 'Monitor answering history' strategy
As shown in Fig 1, FAM-WATA provides an interface to check the answering history of the user and others who have taken the test, available to learners after they pass the test. Through understanding their own progress, learners are expected to take the initiative in monitoring their learning.

Strategy 5: 'Query scores' strategy
FAM-WATA provides an interface for learners to look up peer scores and see the progress of others, to encourage the learner to learn from peers, and motivate learning (Fig 2). Students may find out whether others have passed the test and how many tries are required for others to answer and to pass the test (Fig 2A). Students can query the answering history of other students (Fig 2B). The main purpose of these designs is to add the stimulus of competition. Those who perform well or pass the test will be marked by special signs (like the star marks in Fig 2A), increasing their sense of achievement. In

Item (61.18 %) **A**	Which layer of the Earth's atmosphere contains ozone whose concentration is relatively high? A. thermosphere B. mesosphere C. stratosphere D. troposphere
Options	⊙ (11.06 %) B ⊙ (13.42 %) D ⊙ (13.42 %) A **B** ⊙ (61.18 %) C ⊙ (0.95 %) *<<NO Answer>>*
Concept Code	0201 020101
Answer	C **C**
Hint	You may refer to e-Learning resources: Ch2.1.1 Further internet resources about the Ozone Layer: http://www.tycg.gov.tw/cgi-bin/SM_theme?page=3e7189ba **D**
k10811 Correct Ratio: 60 %	Times 1: B Times 2: D Times 3: C YES!! Times 4: C YES!! **E** Times 5: C YES!! IS (Item Score):6

Fig 1 'Monitor answering history' strategy: (A) Students can know the ratio of correct answers for each item (B), the rate at which each option is selected, (C) the correct answer, (D) references, and (E) personal answering history.

Fig 2A, the student (k10821) has passed the test after seven tries, while a10505 required 10 tries. Because k10821 has passed after only seven tries, his performance in the FAM-WATA is better than that of a10505.

Strategy 6: 'All pass and then reward' strategy
FAM-WATA will generate a FLASH (Adobe Systems Inc., CA, USA) animation to congratulate learners on passing the test. Animation effects can stimulate learner interest (Mayer & Moreno 2002). This type of positive feedback can also be regarded as a form of encouragement for learners who pass a task, creating a sense of achievement.

Evaluating FAM-WATA

The major purpose of this research was to examine the potential benefits of FAM-WATA on student learning. First, this research used a cognitive style approach to investigate and compare the effectiveness of different formative assessment types in e-learning environment. Three different types of formative assessment (TFA) used in e-learning environment were paper-and-pencil test (PPT), normal web-based formative assessment (N-WATA), and FAM-WATA (see Table 2). In addition, this research investigated whether student performance in FAM-WATA was related to student achievement in an e-learning environment.

Participants

Participants for this research comprised 516 seventh-grade students in 14 classes and eight teachers from schools located across central Taiwan. A total of 503 seventh-grade students (260 male and 243 female) were valid for data analysis. All eight teachers had experience in constructing e-learning environments. The 14 classes of the eight teachers were randomly divided into three different groups: the PPT group, the N-WATA group, and the FAM-WATA group (see Table 3). Each class as a unit took the e-learning course (*Atmosphere and Water*) and participated in only one of the three TFA. The seventh graders in this research all took a related course, *Introduction to Computers and Internet*, to familiarize them with computers and the Internet.

A

Student ID	Scores	Test Time	Used Time
k10821 (times: 1)	80	2006-08-09 11:00:36	1.22
k10821 (times: 2)	100	2006-08-09 11:01:37	0.8
k10821 (times: 3)	80	2006-08-09 11:02:41	0.94
k10821 (times: 4)	100	2006-08-09 11:03:15	0.45
k10821 (times: 5)	100	2006-08-09 11:04:06	0.8
k10821 (times: 6)	100	2006-08-09 11:04:37	0.45
★ k10821 (times: 7)	100	2006-08-09 11:05:01	0.37
a10505 (times: 1)	40	2006-08-10 16:34:16	1.44
a10505 (times: 2)	20	2006-08-10 16:36:03	1.62
a10505 (times: 3)	40	2006-08-10 16:37:43	1.35
a10505 (times: 4)	80	2006-08-10 16:40:13	1.05
a10505 (times: 5)	80	2006-08-10 16:41:57	1.42
a10505 (times: 6)	100	2006-08-10 16:42:35	0.5
a10505 (times: 7)	100	2006-08-10 16:42:49	0.19
a10505 (times: 8)	100	2006-08-10 16:43:06	0.25
a10505 (times: 9)	100	2006-08-10 16:43:16	0.12
★ a10505 (times: 10)	100	2006-08-10 16:43:21	0.07

B

Item (71.81 %)	**Which reason makes the sea water taste salty?**
Options	⊙ (16.18 %) The sea water contains magnesium chloride ($MgCl_2$) ⊙ (6.53 %) The sea water contains magnesium carbonate ($MgCO_3$) ⊙ (71.81 %) The sea water contains sodium chloride (NaCl) ⊙ (3.95 %) The sea water contains calcium carbonate ($CaCO_3$) ⊙ (1.55 %) *<<No answer>>*
Concept Code	0201 020102
Answer	The sea water contains sodium chloride (NaCl)
Hint	You may refer to e-Learning resources: Ch2.1.2 Further internet resources about the sea water: http://book.tngs.tn.edu.tw/database/scientieic/content/1991/00110263/0012.htm
a10506 Correct Ratio: 75 %	Times 1: The sea water contains magnesium chloride ($MgCl_2$) Times 2: The sea water contains sodium chloride (NaCl) **YES!!** Times 3: The sea water contains sodium chloride (NaCl) **YES!!** Times 4: The sea water contains sodium chloride (NaCl) **YES!!** **IS (Item Score):** 7.5

Fig 2 'Query scores' strategy: (A) Students can query personal and others scores and whether or not others have passed the test. (B) Students can query others' answering history of each item.

Instruments

WATA Formative Assessment Strategies Scale (WFASS)
The WFASS was used to evaluate student attitude towards the strategies built into FAM-WATA. Six subscales were used to evaluate student attitudes towards the six strategies in FAM-WATA. All subscales used a five-point Likert scale, including 'strongly agree (5 points)', 'agree (4 points)', 'neutral (3 points)', 'disagree (2 points)', and 'strongly disagree (1 point)'. An average above 3.00 points thus represented a positive attitude towards the strategy.

The content of the WFASS was constructed by the author and refined by three assessment experts. The WFASS was administered to 280 junior high school students for pilot-testing and item deletion. The final version of WFASS had 22 items. The Cronbach's α of each subscale was: 'repeat the test' subscale: 0.76; 'correct answers are not given' subscale: 0.70; 'query scores' subscale: 0.71; 'ask questions' subscale: 0.77; 'monitor answering history' subscale: 0.82; and 'all pass and then reward' subscale: 0.79.

Formative assessment and summative assessment
This research developed a Biology e-learning course on the topic of *Atmosphere and Water*. The learning content of the e-learning course was divided into three sections, with one formative assessment used in each section. This research constructed the test items of the formative assessments according to the learning contents of each section. Students in the three groups all practised the same formative assessment items,

Table 2 Comparison of FAM-WATA, N-WATA and PPT.

Strategies	FAM-WATA	N-WATA[a]	PPT
Repeat the test	1. Students may repeat the test for more practice. The system randomly chooses at most five items for students to answer each time. 2. If the students answer a given item correctly three times consecutively, the item will not appear again on the test.	1. Students may repeat the test for more practice. Students have to answer all the items every time. 2. The system does not create an answering history to judge whether students have passed the test or not.	N/A[c]
Correct answers are not given	After students submit the answers, correct answers are not given directly, but a reference is provided.	N/A[b]	N/A[d]
Ask questions	Students are allowed to ask questions of teachers asynchronously through the system.	YES	Students are allowed to ask questions in class.
Query scores	Students may check their own scores and those of peers.	YES	Students may know their own scores and check peer scores privately.
Monitor answering history	Students can query personal and peer answering history of each item.	N/A	N/A
All pass and then reward	When students reach the level of 'pass the test', the system will show an animation as a reward.	N/A	N/A

[a] Normal WATA (N-WATA) is a normal web-based formative assessment design, not equipped with all six FAM-WATA strategies. It uses the WATA system to simulate the normal web-based formative assessment.

[b] The system will provide correct answers directly, along with a reference.

[c] The formative assessment plays a role quite similar to quizzes in class. Students cannot repeat the test in class. Moreover, test papers are handed back to students, and they can practice the questions after class.

[d] Correct answers are announced by the teacher, but the answers are not deliberately explained.

YES, equip the same strategy as the FAM-WATA; N/A, not equip the strategy; FAM-WATA, Formative Assessment Module of the Web-based Assessment and Test Analysis System; N-WATA, normal web-based formative assessment; PPT, paper-and-pencil test.

Table 3. Students of the three different groups by cognitive style.

Cognitive Style	Group	Gender		Sum	
Field independent	PPT	Female	44	84	256
		Male	40		
	N-WATA	Female	45	101	
		Male	56		
	FAM-WATA	Female	31	71	
		Male	40		
Field dependent	PPT	Female	26	65	247
		Male	39		
	N-WATA	Female	42	77	
		Male	35		
	FAM-WATA	Female	55	105	
		Male	50		

FAM-WATA, Formative Assessment Module of the Web-based Assessment and Test Analysis System; N-WATA, normal web-based formative assessment; PPT, paper-and-pencil test.

differing only in the administering approach of the formative assessments (see Table 2).

For the summative assessment, however, this research constructed test items according to the learning contents of the whole e-learning course to evaluate learning effectiveness. Pretest scores of the summative assessment were taken to represent entry behaviour of learning, while the difference between post-test scores and pretest scores was taken to represent learning effectiveness. There were 45 items in the summative assessment. The average difficulty index for the summative assessment was 0.52. The Cronbach's α for the summative assessment was 0.90. The test items in the formative assessment were not repeated in the summative assessment.

Hidden Figures Test (HFT)
The cognitive style dimension investigated in this research is Witkin's 'field dependent/field independent'. There are many instruments developed to determine individual relative 'field dependent/field independent' learning behaviour, including the Group Embedded Figures Test/Embedded Figures Test (GEFT/EFT) (see Witkin *et al*. 1971), the HFT (Educational Testing Service 1962; French *et al*. 1963) and the Portable Rod and Frame Test (PRFT) (Oltman 1968). This research used the HFT as the instrument to investigate the cognitive style of individual learners.

In the HFT, participants are presented with five simple figures hidden inside more complex figures. Participants must identify the simple figures. There are 32 such complex figures in the HFT, divided into two sections, each of which consists of 16 complex figures. Each section was administered separately. Both correctly and incorrectly answered items were recorded to calculate the final scores. The final scores of each participant were used to determine his/her cognitive style.

With regard to the reliability of the HFT, Boersma (1968) found that the test–retest reliability index of the HFT was 0.63. In addition, Shapiro (1970) found that the split-half reliability index of the HFT was between 0.58 and 0.80. As for validity, Shapiro (1970) found that the correlation index between HFT and EFT was 0.51.

Procedures

To lower the influence of different learning contents as well as teacher and student familiarity on the results of this research, the eight teachers were asked to join seminars and teaching presentations so that they could understand the purpose of this research and teaching methods in advance. In addition, before students participated in this research, the eight teachers had lectured over the Web to enable their students to become accustomed to the e-learning environment.

Before the start of the e-learning course, students who enrolled in each formative assessment group were given the pretest separately. This research required 2 weeks in total (six classes). During the six classes, students in the FAM-WATA and N-WATA groups could learn on the Web and take part in the web-based

formative assessments at any time. Students in the PPT group could also study on the Web at any time, but formative assessments (paper-and-pencil test) were administered after each class. Correct answers were given but not actively explained by teachers. Instead, students were supposed to actively ask teachers questions and were asked to find answers from the web pages. Furthermore, test papers were returned to students to be used for review. After being taught for 2 weeks, all the students had to take the post-test. Additionally, the FAM-WATA group had to finish the WFASS.

Data collection and analyses

The data collected in this research were all quantitative data, including the pretest and post-test scores of the summative assessment, WFASS scores, and accumulated scores (AS). Upon completion of the data gathering phase, the scores were analysed with SPSS™ Version 10.0 (SPSS Inc, Chicago, IL, USA). Because the six subscales in the WFASS used a five-point Likert scale, an average score of each subscale above 3.00 points represented a positive attitude towards the strategy. Internal consistency (Cronbach's α) was also analysed.

This research used analysis of covariance (ANCOVA), taking 'pretest scores of summative assessment (PRE)' as the covariate and 'post-test scores of summative assessment (POST)' as the dependent variable, to test the relationship between the POST and the TFA. There were three different TFA: FAM-WATA, N-WATA, and PPT. Least significant difference (LSD) was used to test the differences among the three TFA.

This research also investigated the learning effectiveness of students with different cognitive styles in three different TFA separately. The data was divided into two parts by participant cognitive style, and was tested by ANCOVA separately. This research used ANCOVA with PRE as the covariate and POST as the dependent variable to test the relationship between the POST and the TFA. There were also three different TFA: FAM-WATA, N-WATA, and PPT. LSD was used to test the differences among the three TFA in two different cognitive style groups separately.

In addition, in order to investigate whether field independent students or field dependent students made more effective use of FAM-WATA, this research also used ANCOVA with PRE as the covariate and POST as the dependent variable to test the relationship between the cognitive style and the POST in the FAM-WATA group.

Finally, the AS was analysed. The AS is a special score calculated based on student performance in FAM-WATA. The formula for the AS of each student in each FAM-WATA assessment is

$$\textbf{Item Score (IS)} \begin{cases} \textbf{Failed Item} \\ (\text{Item NOT correctly answered} \\ \text{three times consecutively}): \textbf{IS} = 0 \\ \textbf{Passed Item} \\ (\text{Item correctly answered in} \\ \text{three times consecutively}): \\ \textbf{IS} = (3/Times)*(100/Total) \end{cases}$$

$Times$: Number of times each
item is totally answered
$Total$: Total number of items in
the FAM-WATA assessment

$$\textbf{AS (Accumulated Score)} = \sum_{n=1}^{Total} ISn$$

$Total$: Total number of items in
the FAM-WATA assessment

The ANCOVA, using PRE as the covariate and POST as the dependent variable, was used to investigate the relationship between student performance in FAM-WATA (AS) and student achievement. During ANCOVA analysis, student average ASs were ranked into three groups, upper, middle, and lower, before data analysis. The upper group comprised students with an average AS in the upper 33% of all scores, while the middle and lower groups represent the middle and lower third of scores respectively. These three groups were defined as fixed factor – ranks of average accumulated scores (RAAS). In addition, this research also used LSD to test the differences among three groups of the RAAS.

Results

Junior high school student attitude towards FAM-WATA

There were 176 valid participants in the FAM-WATA group. The results indicated that the Cronbach's α of almost all subscales was over 0.70 and the average scores of all subscales in the WFASS were all above 3.00 points (see Table 4), showing that the students held positive attitudes towards the six strategies of the FAM-WATA.

Table 4. Student attitude towards FAM-WATA ($n = 176$).

Subscale	Items	Average	SD	Cronbach's α
Repeat the test	6	4.06	0.31	0.76
Correct answers are not given	3	4.10	0.04	0.72
Query scores	4	3.99	0.11	0.73
Ask questions	3	4.23	0.10	0.76
Monitor answering history	3	4.05	0.21	0.70
All pass and then reward	3	4.22	0.02	0.78

Learning effectiveness of three different TFA

Before ANCOVA, the homogeneity of variance assumption was tested. The Levene's test for equality of variances was not significant ($F = 0.23$, $P > 0.05$), indicating the variances are homogeneous and the homogeneity assumption had been met. Table 5 shows that the PRE had a significant impact on the POST ($F = 193.06$, $P < 0.01$). Furthermore, the TFA was found to have a significant impact on the POST ($F = 4.55$, $P < 0.05$) (see Table 5).

Table 5 also shows that the FAM-WATA group performed significantly better than the N-WATA group and PPT group ($P < 0.05$). In addition, the N-WATA group performed better than the PPT group, although not significantly so. In summary, different TFA appear to affect e-learning effectiveness, and FAM-WATA appears to be more effective.

Learning effectiveness of students with different cognitive styles in three TFA

There were 256 valid participants in the field independent group and 247 valid participants in the field dependent group, for a total of 503 valid participants. Before ANCOVA, the homogeneity of variance assumption of each cognitive style group was tested. The Levene's test for equality of variances was not significant (field independent group: $F = 0.09$, $P > 0.05$; field dependent group: $F = 0.49$, $P > 0.05$) for both groups, indicating the variances are homogeneous and the homogeneity assumption had been met in both groups. Table 6 shows that the PRE had a significant impact on the POST in both group (field independent group: $F = 81.53$, $P < 0.01$; field dependent group: $F = 98.77$, $P < 0.01$). However, the TFA had a significant impact on the POST only in the field independent group ($F = 4.43$, $P < 0.05$).

Table 6 also shows that field independent students in the FAM-WATA group performed significantly better than field independent students in the N-WATA group and PPT group ($P < 0.05$). In addition, field independent students in the N-WATA group also performed better than field independent students in the PPT group, although not significantly so. With regard to the field dependent group, field dependent students in the FAM-WATA group performed better than field dependent students in the N-WATA group and PPT group, although not significantly so. Field dependent students in the N-WATA group also performed better than field dependent students in the PPT group, although not significantly so.

Table 5. Descriptive statistics and ANCOVA ($n = 503$).

Variable	Level	Mean (SD)	F	Post hoc
PRE	FAM-WATA	40.40 (13.39)	193.06**	N/A
	N-WATA	40.90 (12.65)		
	PPT	45.06 (14.33)		
TFA	FAM-WATA	54.65 (18.61)	4.55*	FAM-WATA > PPT*
	N-WATA	50.83 (18.16)		FAM-WATA > N-WATA*
	PPT	53.31 (19.06)		

*$P < 0.05$; **$P < 0.01$.
FAM-WATA, Formative Assessment Module of the Web-based Assessment and Test Analysis System; N-WATA, normal web-based formative assessment; PPT, paper-and-pencil test; PRE, pretest scores; TFA, types of formative assessment.

Table 6. Descriptive statistics and ANCOVA of cognitive styles.

Cognitive style	Variable	Level	Mean (SD)	F	Post hoc
Field independent (n = 256)	PRE	FAM-WATA	41.58 (13.22)	81.53**	N/A
		N-WATA	42.48 (11.96)		
		PPT	47.90 (13.97)		
	TFA	FAM-WATA	59.09 (17.22)	4.43*	FAM-WATA > PPT*
		N-WATA	53.65 (17.40)		FAM-WATA > N-WATA*
		PPT	56.46 (18.53)		
Field dependent (n = 247)	PRE	FAM-WATA	39.59 (13.51)	98.77**	N/A
		N-WATA	38.84 (13.29)		
		PPT	41.39 (14.06)		
	TFA	FAM-WATA	51.66 (19.00)	1.78	
		N-WATA	47.12 (18.56)		
		PPT	49.23 (19.11)		

*$P < 0.05$; **$P < 0.01$.
FAM-WATA, Formative Assessment Module of the Web-based Assessment and Test Analysis System; N-WATA, normal web-based formative assessment; PPT, paper-and-pencil test; PRE, pretest scores; TFA, types of formative assessment.

Table 7. Descriptive statistics and ANCOVA table (n = 176).

Variable	Level	Mean (SD)	F	Post hoc
PRE	Field Independent	41.58 (13.22)	81.10**	N/A
	Field dependent	39.59 (13.51)		
Cognitive style	Field Independent	59.09 (17.22)	6.38*	Field Independent > Field dependent*
	Field dependent	51.66 (19.00)		

*$P < 0.05$; **$P < 0.01$.
PRE, pretest scores.

To investigate students with which cognitive style enabled better performance in the FAM-WATA group, ANCOVA was used to analyse the data from the FAM-WATA group. There were 176 valid participants in the FAM-WATA group, including 105 field dependent students and 71 field independent students. Before ANCOVA, the homogeneity of variance assumption was also tested. The Levene's test for equality of variances was not significant ($F = 0.01$, $P > 0.05$), representing that the variances are homogeneous and the homogeneity assumption had been met. Table 7 shows that the PRE had a significant impact on the POST ($F = 81.10$, $P < 0.01$). The cognitive style also had a significant impact on the POST ($F = 6.38$, $P < 0.05$). Table 7 also shows that field independent students performed significantly better than field dependent students in the FAM-WATA group ($P < 0.05$).

In summary, by comparison with N-WATA and PPT, field independent students appeared to get the most out of the e-learning environment equipped with FAM-WATA. However, field dependent students appeared to have equal achievement in the e-learning environment regardless of the TFA. In addition, field independent students performed significantly better than field dependent students in the e-learning environment equipped with FAM-WATA. Thus, FAM-WATA was more effective in facilitating the e-learning effectiveness of field independent students.

Relationship between student performance in FAM-WATA and their learning effectiveness
This research used the ANCOVA to investigate whether student performance in FAM-WATA, AS, was related to student learning effectiveness in an e-learning environment.

Before ANCOVA, the homogeneity of variance assumption was tested. The Levene's test for equality of variances was not significant ($F = 0.19$, $P > 0.05$),

Table 8. Descriptive statistics and ANCOVA (n = 176).

Variable	Level	Mean (SD)	F	Post hoc
PRE	Upper group	44.31 (12.97)	76.10**	N/A
	Middle group	37.29 (11.47)		
	Lower group	39.57 (14.80)		
RAAS	Upper group	62.07 (18.31)	6.65**	Upper group > Lower group**
	Middle group	53.27 (16.07)		Middle group > Lower group*
	Lower group	48.52 (19.04)		

*$P < 0.05$; **$P < 0.01$.
PRE, pretest scores; RAAS, ranks of average accumulated scores.

indicating the variances are homogeneous and the homogeneity assumption had been met. Table 8 shows that the PRE had a significant impact on the POST (F = 76.10, $P < 0.01$). In addition, the RAAS had a significant impact on the POST (F = 6.65, $P < 0.01$). The results show that student performance in the FAM-WATA had a significant impact on the learning effectiveness.

Table 8 also shows that the upper group performed significantly better than the lower group ($P < 0.01$). The middle group performed significantly better than the lower group ($P < 0.05$). Additionally, the upper group performed better than the middle group, although not significantly so. These results reveal that AS could effectively represent student e-learning effectiveness. All in all, the higher the student AS in FAM-WATA, the better the student achievement in the e-learning environment.

Concluding remarks

Many researchers have emphasized the importance of formative assessment in student learning achievement (Brown & Knight 1994; Boud 1995; Barron *et al*. 1998; Black & Wiliam 1998; Bransford *et al*. 2000, pp. 140–141; Buchanan 2000; Khan *et al*. 2001; Manogue *et al*. 2002; Velan *et al*. 2002; Henly 2003), but studies on formative assessment strategy and its effects are not plentiful (Black & Wiliam 1998; Bransford *et al*. 2000; Bell & Cowie 2001; Velan *et al*. 2002; Henly 2003; Wang *et al*. 2004a). Most web-based formative assessment strategies found in the literature are akin to those Buchanan (2000) suggested, including 'repeat the test', 'correct answers are not given', and 'timely feedback' (i.e. referring to the related materials when answering). Buchanan (2000) suggested that these strategies greatly benefit learning effectiveness among college students. This research supports those findings.

This research augmented Buchanan's three strategies (Buchanan 2000) with three new strategies: 'query scores', 'monitor answering history', and 'all pass and then reward'. This research adopted a quasi-experimental design to examine the effectiveness of three different TFA in an e-learning environment. The results revealed that embedding web-based formative assessment in an e-learning environment (e.g. FAM-WATA and N-WATA) was better than using paper-and-pencil formative assessment. Moreover, web-based formative assessment with FAM-WATA strategies (FAM-WATA) was significantly better than web-based formative assessment without FAM-WATA strategies (N-WATA) and paper-and-pencil formative assessment (PPT). In other words, in the e-learning environment, learning effectiveness will be enhanced if traditional paper-and-pencil test are replaced by web-based formative assessment. Additionally, the results appeared to show that if a web-based assessment can be equipped with additional instructional strategies such as those designed into FAM-WATA, learning effectiveness in the e-learning environment will be significantly enhanced. The possible reason is that FAM-WATA provides students with a well-designed mechanism to revise the mistakes they have made. However, the factors making FAM-WATA successful need to be further investigated. In the future, the oncoming researches are going to investigate the mechanism by which the FAM-WATA strategies affect e-learning effectiveness.

In addition, the effectiveness of students with different cognitive styles in the three different TFA was also investigated. It was found that only field independent students had significantly different performance among three different TFA. The results also indicated that field independent students in the FAM-WATA group performed significantly better than those in the N-WATA

and PPT group. Moreover, the results revealed that field independent students performed significantly better than field dependent students in the FAM-WATA group. These findings were consistent with the characteristics of field independent/dependent individuals stated by Frank and Keane (1993) and Tinajero and Paramo (1998). Both Frank and Keane and Tinajero and Paramo concluded that field dependent individuals, by comparison with field independent individuals, preferred the strategy of rehearsal (passive cognitive strategy). In this research, the major purpose of the six strategies of FAM-WATA was to make students participate in the web-based formative assessment effectively and spontaneously. However, the FAM-WATA did not directly provide correct answers. Field dependent students thus could not get correct answers to rehearse directly and were likely to have trouble in making use of the FAM-WATA to facilitate their learning. The FAM-WATA appeared not to be suitable for field dependent students. By contrast, the FAM-WATA was more effective for field independent students. The possible reason is that field independent students were likely to take advantage of the FAM-WATA strategies to adopt more active cognitive strategy to categorize and discover the correct answers. Nevertheless, it requires further investigation. From the viewpoint of individualized e-learning design, the FAM-WATA is suggested to be incorporated into the e-learning environment for field independent students.

This research further explored the relationship between student performance in FAM-WATA and their learning effectiveness. The results showed that student performance in FAM-WATA had a significant impact on student learning effectiveness. The results also showed that the better students performed in FAM-WATA, the better students performed in the e-learning environment.

All in all, the strategies of FAM-WATA were successful in facilitating student learning, especially for the field independent students. This research suggests that web-based formative assessment incorporating effective strategies is necessary in a successful e-learning environment. The strategies of the web-based formative assessment in this research, 'repeat the test', 'correct answers are not given', 'query scores', 'ask questions', 'monitor answering history', and 'all pass and then reward', appear to be successful in improving student e-learning. This research also suggests that all these strategies should be used simultaneously. Furthermore,

FAM-WATA, combined with the equation for calculating AS, is also suggested. This research suggests that the AS provides an effective feedback mechanism for students to monitor their formative learning and for teachers to understand the learning condition about their students.

Just as Bell and Cowie (2001), Black and Wiliam (1998), and Bransford *et al.* (2000) have suggested, new formative assessment strategies should be designed and implemented on an ongoing basis. E-learning researchers and innovators who incorporate IT into their instruction should become familiar with strategies of web-based formative assessment and understand clearly the relationship between web-based formative assessment strategies and e-learning effectiveness. Further research into innovative strategies of web-based formative assessment is necessary to help understand how best to use the web-based formative assessment strategies in an e-learning environment.

Acknowledgements

References

Barron B.J.S., Schwartz D.L., Vye N.J., Moore A., Petrosino A., Zech L., Bransford J. & The Cognition and Technology Group at Vanderbilt (1998) Doing with understanding: lessons from research on problem- and project-based learning. *Journal of the Learning Sciences* **7**, 271–311.

Bell B. & Cowie B. (2001) The characteristics of formative assessment in science education. *Science Education* **85**, 536–553.

Black P. & Wiliam D. (1998) Assessment and classroom learning. *Assessment in Education* **5**, 7–74.

Boersma F.J. (1968) Test-retest reliability of the cf-1 Hidden Figures Test. *Educational & Psychological Measurement* **28**, 555–559.

Boud D. (1995) *Enhancing Learning through Self-Assessment*. Routledge, London.

Bransford J.D., Brown A. & Cocking R. (2000) *How People Learn: Mind, Brain, Experience and School, Expanded Edition*. National Academy Press, Washington, DC.

Brewer C.A. (2004) Near real-time assessment of student learning and understanding in Biology courses. *Bioscience* **54**, 1034–1039.

Brown S. & Knight P. (1994) *Assessing Learners in Higher Education*. Kogan Page, London.

Buchanan T. (1998) Using the World Wide Web for formative assessment. *Journal of Educational Technology Systems* **27**, 71–79.

Buchanan T. (2000) The efficacy of a World-Wide Web mediated formative assessment. *Journal of Computer Assisted Learning* **16**, 193–200.

Chen S.Y. & Macredie R.D. (2004) Cognitive modelling of student learning in web-based instructional programmes. *International Journal of Human–Computer Interaction* **17**, 375–402.

Chen S.Y., Magoulas G.D. & Dimakopoulos D. (2005) A flexible interface design for web directories to accommodate different cognitive styles. *Journal of the American Society for Information Science and Technology* **56**, 70–83.

DeTure M. (2004) Cognitive style and self-efficacy: predicting student success in online distance education. *The American Journal of Distance Education* **18**, 21–38.

Ebel R.L. & Frisbie D.A. (1991) *Essentials of Educational Measurement*, 5th edn. Prentice Hall, Inc, Englewood Cliffs, NJ.

Educational Testing Service (1962) *Hidden Figures Test – Cf-1*. Educational Testing Service, Princeton, NJ.

Frank B.M. & Keane D. (1993) The effect of learner's field independence, cognitive strategy instruction, and inherent word-list organisation on free-recall memory and strategy use. *Journal of Experimental Education* **62**, 14–25.

French J.W., Ekstrom R.B. & Price L.A. (1963) *Kit of Reference Tests for Cognitive Factors.* Educational Testing Service, Princeton, NJ.

Gardner L., Sheridan D. & White D. (2002) A web-based learning and assessment system to support flexible education. *Journal of Computer Assisted Learning* **18**, 125–136.

Ghinea G. & Chen S.Y. (2003) The impact of cognitive styles on perceptual distributed multimedia quality. *British Journal of Educational Technology* **34**, 393–406.

Gipps C. (1994) *Beyond Testing: Towards a Theory of Educational Assessment*. The Falmer Press, London.

Goodenough D. (1976) The role of individual differences in field dependence as a factor in learning and memory. *Psychological Bulletin* **83**, 675–694.

Harrison A.W. & Rainer R.K. (1992) The influence of individual differences on skill in end-user computing. *Journal of Management Information Systems* **9**, 93–111.

Henly D.C. (2003) Use of Web-based formative assessment to support student learning in a metabolism/nutrition unit. *European Journal of Dental Education* **7**, 116–122.

Jonassen D.H. & Grabowski B. (1993) *Individual Differences and Instruction.* Allyn & Bacon, New York.

Justham D. & Timmons S. (2005) An evaluation of using a web-based statistics test to teach statistics to post-registration nursing students. *Nurse Education Today* **25**, 156–163.

Khan K.S., Davies D.A. & Gupta J.K. (2001) Formative self-assessment using multiple true-false questions on the Internet: feedback according to confidence about correct knowledge. *Medical Teacher* **23**, 158–163.

Kim K.-S. (2001) Implications of user characteristics in information seeking on the World Wide Web. *International Journal of Human–Computer Interaction* **13**, 323–340.

Manogue M., Kelly M., Masaryk S.B., Brown G., Catalanotto F., Delap E., Godoroja P. & Morio I. (2002) Evolving methods of assessment. *European Journal of Dental Education* **6**, 53–66.

Mayer R.E. & Moreno R. (2002) Animation as an aid to multimedia learning. *Educational Psychology Review* **14**, 87–99.

Oltman P.K. (1968) A portable rod-and-frame apparatus. *Perceptual and Motor Skills* **26**, 503–506.

Peat M. & Franklin S. (2002) Supporting student learning: the use of computer-based formative assessment modules. *British Journal of Educational Technology* **33**, 515–523.

Perrenoud P. (1998) From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. *Assessment in Education* **5**, 85–102.

Sadler D.R. (1989) Formative assessment and the design of instructional systems. *Instructional Science* **18**, 119–144.

Scriven M. (1967) The methodology of evaluation. In *Perspectives of Curriculum Evaluation* (ed. R. Tyler), pp. 39–83. AERA Monograph Series on Curriculum Evaluation (no. 1). Rand McNally, Skokie, IL.

Shapiro N.P. (1970) *The effects of perceptual isolation on recall for field-independent and field-dependent subjects*. Unpublished Doctoral Dissertation, New York University, New York.

Sujo de Montes L.E. & Gonzales C.L. (2000) Been there, done that: reaching teachers through distance education. *Journal of Technology and Teacher Education* **8**, 351–371.

Tinajero C. & Paramo M.F. (1998) Field dependence-independence and strategic learning. *International Journal of Educational Research* **29**, 251–262.

Velan G.M., Kumar R.K., Dziegielewski M. & Wakefield D. (2002) Web-based self-assessments in pathology

with Questionmark Perception. *Pathology* **34**, 282–284.

Wang T.H., Wang K.H., Wang W.L., Huang. S.C. & Chen S.Y. (2004a) Web-based assessment and test analyses system: development and evaluation. *Journal of Computer Assisted Learning* **20**, 59–71.

Wang T.H., Wang K.H., Wang W.L. & Huang, S.C. (2004b) Assessing effectiveness of formative assessment models on students' learning achievement in e-learning environment. *Chinese Journal of Science Education* **12**, 469–490.

Wiliam D. & Black P. (1996) Meanings and consequences: a basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal* **22**, 537–548.

Williams C.E. (2004) School district of Philadelphia uses Web-based system to increase student achievement. *T H E Journal* **31**, 51–52.

Witkin H.A. (1950) Individual differences in ease of perception of embedded figures. *Journal of Personality* **19**, 1–15.

Witkin H. & Asch S.E. (1948) Studies in space orientation: III. Perception of the upright in the absence of a visual field. *Journal of Experimental Psychology* **38**, 603–614.

Witkin H.A., Moore C.A., Goodenough D.R. & Cox P. (1977) Field dependent and field independent cognitive styles and their educational implications. *Review of Educational Research* **47**, 1–64.

Witkin H.A., Oltman P.K., Raskin E. & Karp S.A. (1971) *A Manual for the Group Embedded Figures Test*. Consulting Psychologists Press, Palo Alto, CA.

Zoe L.R. & DiMartino D. (2000) Cultural diversity and end-user searching: an analysis by gender and language background. *Research Strategies* **17**, 291–305.