Computers and Education: Artificial Intelligence 2 (2021) 100015

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.elsevier.com/locate/caeai

An interactive test dashboard with diagnosis and feedback mechanisms to facilitate learning performance

Chih-Ming Chen [a,*,1], Jung-Ying Wang [b,2], Li-Chieh Hsu [a,3]

[a] *Graduate Institute of Library, Information and Archival Studies, National Chengchi University, No. 64, Section 2, ZhiNan Road, Wenshan District, Taipei City, 116, Taiwan, ROC*

[b] *Department of Multimedia and Game Science, Lunghwa University of Science and Technology, No.300, Section1, Wanshou Rd., Guishan District, Taoyuan City, 333326, Taiwan, ROC*

A R T I C L E I N F O

Keywords:
Digital dashboard for learning (DDL)
Test dashboard with diagnosis and feedback
Computer-adaptive science assessment
Physics self-efficacy
Technology acceptance

A B S T R A C T

This study developed an interactive test dashboard with diagnosis and feedback mechanisms (ITD-DFM) that can generate visualized, rich, and high-quality test feedback for each learner's learning reflection and review based on simultaneously considering test response time and correctness to assist students' learning. A quasi-experimental research was conducted with 50 Grade 8 students from two classes in a public junior high school in Taiwan to assess the learning performance of ITD-DFM. One class with 26 students was assigned to the experimental group using the ITD-DFM to support learning, whereas the other class with 24 students was assigned to the control group using the traditional test system without diagnosis and feedback mechanisms (TTS-NDFM). Experimental results reveal that the learning performance, physics self-efficacy, and technology acceptance of the experimental group were significantly better than those of the control group. The ITD-DFM has the same effect on promoting the learning performance of learners with different prior knowledge levels as well as learners with either high or low prior knowledge level in the experimental group exhibited significant improvement in physics self-efficacy, but no such results in the control group. The ITD-DFM provides more benefit in promoting the technology acceptance of learners with high prior knowledge level.

1. Introduction

Learning assessment plays an essential role in the learning processes of students. In educational settings, learning assessment is generally divided into three types: assessment for learning, assessment of learning, and assessment as learning (Earl & Katz, 2006). Assessment for learning or called as formative assessment involves continuous assessment and feedback given by teacher or real-time learning process analysis to understand students' learning progress and outcomes during learning processes (Chen, Chen, & Horng, 2019; Nicol & Macfarlane-Dick, 2006). Assessment of learning or called as summative assessment refers to the assessment of learning progress and outcomes of students within a definite time as well as it consists of testing or accumulating evidence regarding to each student's learning outcomes over time at the end phase of a learning process (Knight, 2002). Assessment as learning or called self-assessment involves critical analysis of oneself to identify the strength and weaknesses and thereby strive for achieving learning progress and goal (Earl, 2003). From the perspective of assisting students' learning, each type of learning assessment has its values and meanings.

The test is the most common way of summative assessment used to assess learning outcomes of students in educational scenarios because it not only helps students understand their current learning situations and problems but also provides them learning feedback to facilitate learning reflection and review. Summative assessment in classrooms is typically done at the end of something (eg, a unit, a course, or a program) and

* Corresponding author.
E-mail addresses: chencm@nccu.edu.tw (C.-M. Chen), wyy@mail.lhu.edu.tw (J.-Y. Wang), chutinghsui@gmail.com (L.-C. Hsu).

[1] **Chih-Ming Chen** is currently a professor in the Graduate Institute of Library, Information and Archival Studies at National Chengchi University, Taipei, Taiwan. His research interests include digital library, e-learning, data mining, machine learning and intelligent agents on the web.

[2] **Jung-Ying Wang** is currently a professor in the Department of Multimedia and Game Science at Lunghwa University of Science and Technology, Taoyuan City, Taiwan. His research interests include game AI, machine learning and intelligent systems on the web.

[3] **Li-Chieh Hsu** is currently a graduate student in the Graduate Institute of Library, Information and Archival Studies at National Chengchi University, Taipei, Taiwan. Her research interests include e-learning and learning assessment.

https://doi.org/10.1016/j.caeai.2021.100015
Received 10 November 2020; Received in revised form 20 February 2021; Accepted 20 February 2021

takes the form of tests or exams that include questions drawn from the material studied during learning time. However, the traditional test generally only provides the number of questions answered correctly or incorrectly, the test scores, and their ranking in the class as the main feedback signals for individual learners' review. In this manner, learners are unable to obtain useful and enough feedback and reflection information to diagnose their learning obstacles and difficulties through these kinds of feedback signals from the traditional test as well as understand where their strengths and weaknesses lie (Vie, Popineau, Bruillard, & Bourda, 2017). A more useful way to promote students' learning performance based on feedback signals from test results is that students can immediately receive the test feedback toward more perceptive method such as visualized information about the concepts error and the teacher can be able to obtain real-time information about the student's interaction with the assessment materials. Hence in recent years, many digital dashboards for learning (DDL) have been developed to assist teachers and students to monitor and reflect their online teaching and learning behavior patterns in virtual learning environments (Riofrío-Luzcando, Diego, et al., 2019; Nyland, 2018; Schwendimann et al., 2017). The DDL is an assisted learning tool that provides visualized feedback to facilitate learning performance based on mining or analyzing students' learning processes. Namely, the main objective of DDL is to improve learning and teaching performance. Therefore, DDL has gradually become an important feedback tool that can support teaching and learning analysis. Schwendimann et al. (2017) reviewed 55 papers of DDL and categorized DDL's functions into three groups, including self-monitoring, monitoring others, and administrative monitoring. Additionally, the indicators used in DDL can be categorized into six broad groups, including learner-related indicators that present information describing the learner(s), action-related indicators that present information about the actions performed by the learner(s), content-related indicators that provide information about the content that the learner(s) interacted with or produced, result-related indicators that give information about the outcome of learners' activities, context-related indicators that provide information about the context where the learning took place, and social-related indicators that show how learners interacted with others (Schwendimann et al., 2017). To the best of our knowledge and literature review, no studies have paid attention to developing test feedback dashboards with indicators designed for visualized test feedback. To fill the research gap, this work thus developed an ITD-DFM in which designs several functionalities to support visualized test feedback related to the correctness of the answer and the response time of answering test items to provide an effectively assisted learning tool to improve the learning performance of individual learners. The developed ITD-DFM intends to extend the role of summative assessment for learning by assisting learners to monitor what the strengths and weaknesses are in their learning processes and use the feedback from this monitoring to make adjustments, adaptations, and even major changes in what they understand. Restated, the developed ITD-DFM can be regarded as a novel learning assessment tool simultaneously considering summative assessment and self-assessment.

To assess the effects of using the developed ITD-DFM to support the learning activity of a physics course, this work examines whether the learning performance, physics self-efficacy, and technology acceptance of the learners who use the developed ITD-DFM to support the learning activity in a physics course are significantly better than those of the learners who use the TTS-NDFM. Additionally, several previous studies (Chen, Chen et al., 2019; Chen, Wang, & Lin, 2019; Liu, Andre, & Greenbowe, 2008) indicated that computer-supported learning systems generally provide more benefits in terms of promoting learning performance for the learners with low prior knowledge in comparison with the learners with high prior knowledge. Thus, whether the learning performance, physics self-efficacy, and technology acceptance of the learners with different prior knowledge levels using the developed ITD-DFM to support the learning activity in a physics course are significantly better than those of the learners with different prior knowledge levels using the TTS-NDFM was also examined in this study.

2. Literature review

2.1. Design and applications of learning dashboard with feedback

The learning dashboard is derived from the design concepts of aircraft dashboards and car dashboards that simultaneously use the numerical values, colors, and graphics to reflect the current flight or driving conditions (Park & Jo, 2015). Using the dashboard's concepts from aircraft and car, the learning dashboard transforms the quantitative data of the learning portfolios and the learning results into visualized information for individual learners. Providing individual learners to understand their learning situations based on digital learning dashboard can help learners control, track, and reflect on their learning situations in e-learning environments (Bajzek, Brown, Lovett, & Rule, 2007; Park & Jo, 2015; Verbert et al., 2013, 2014). Besides, the learning portfolios of the learners displayed in the learning dashboard can also be provided to educators as a reference to understand the learning process and progress of each learner. Therefore, the learning dashboard can be used as a learner's self-monitoring tool or as a supervisor to supervise her/his learning processes (Schwendimann et al., 2017; Yu, 2017). The visualized presentation can transform abstract data into visualized graphics and let learners more quickly understand the important information than text-based information. In other words, visualized feedback is one of the most direct and effective ways to transfer and process knowledge and it is the main feature of the learning dashboard (Ellis, 2003; Schwendimann et al., 2017).

Also, many recent studies have paid attention to conducting the literature reviews of the learning dashboard, which help new researchers to acquire knowledge about how to design and develop learning dashboard. For example, Jivet, Scheffel, Specht, and Drachsler (2018) and Schwendimann et al. (2017) reviewed 26 and 55 papers of learning analytics dashboard, respectively, indicating that teachers and students will be able to obtain real-time information about how, where, and when to study from learning analytics dashboard. Park and Jo (Park & Jo, 2019) also reviewed many previous studies on the learning analytics dashboard to suggest the need of developing a tool to assess the effectiveness of the learning analytics dashboard. Schwendimann et al. (2017) proposed a systematic literature review on learning dashboards in the fields of learning analytics and educational data mining. However, based on the authors' best knowledge, no existing studies focused on developing a test dashboard. This work thus designed an ITD-DFM that can simultaneously consider the test response time and correctness as the review signals to provide the visualized test feedback information for individual learners to promote their learning performance.

2.2. Digital dashboards design for promoting learning performance

A digital dashboard for learning (DDL) aims to support the learner in the learning process and the teacher in the teaching process by summarizing and visualizing various information. To implement a DDL, there are five key points proposed by Bajzek et al. (2007) that have to be considered. They include (1) Construction of interactive learning objects; (2) Tagging of all materials with key learning indicators and keeping learning units tied to articulated learning objectives; (3) Collection of data by the dashboard server; (4) Real-time data mining and analysis of data; (5) Construction of dashboard to display information intuitively and allow the learners and instructors to interact with this information. Developing useful and timely feedback on a DDL system for students and teachers is a very challenging task. Hence, some studies investigated and compared the design of various DDLs. For example, Verbert et al. (2013) compared 15 dashboard applications in supporting learning, indicating that most of these dashboards rely on a variety of data gathered from the learning environment. Meanwhile, most of the systems keep track of the time spent (Nine of 15). Moreover, de Freitas et al. (2017) used gamified dashboards to provide immediate feedback of tracking learning performance for university students. Their study focused on how data can be appropriately visualized and presented on DDL to effectively support students' learning. Particularly, the gamified elements, such as points, leveling up, narrative, and progression are used as scaffoldings in a gamified dashboard. Analytical results confirmed that a gamified dashboard is a useful tool to increase students' learning motivation, learning engagement, learning satisfaction, learning retention, and learning performance.

Based on the above literature review, this work refers to the visualization model of DDL proposed by Verbert et al. (2013) and Schwendimann et al. (2017), that simultaneously considered the four elements including awareness, reflection, sense-making, and behavioral change to develop an ITD-DFM. Especially, the proposed ITD-DFM also considers visualizing the time spent in answering each test question because the reaction time in a test, which is known as the interval of time between receiving a question and activating a response, can show whether a learner is familiar with the concepts he/she learned or not (Magill, 2007).

3. The proposed ITD-DFM

The proposed ITD-DFM was implemented by HTML, PHP, JavaScript programming languages, and MySQL database. Particularly, the visualized functionalities developed in the ITD-DFM were implemented by using the suite of chart.js in JavaScript. This section aims to describe the user interface and functionalities of the proposed ITD-DFM.

3.1. The test interface of the ITD-DFM

Fig. 1 shows the test interface of the ITD-DFM. In Fig. 1, the number of questions is displayed on the left side, and there are a total of 30 questions in this case. The red color represents the question that is being answered, and the green color indicates the question number that has been answered. The middle part of the test interface shows the content of the question that is being answered, including the title description, icon, and four answer choices. After all the questions have been answered, users can press the icon in the lower-left corner to submit the test. The time for answering all the questions was limited to 20 min. At the ending of the test time, the system will automatically stop the test process and send out the student's answer to the system. This study selected the learning unit of light in the physics textbook of Taiwan's junior high school as the learning materials. The learning unit of light can be further divided into 5 sections, which are: (1) light propagation, (2) light reflection and mirror, (3) light refraction and lens, (4) optical instrument, and (5) light and color. In this study, each section has six corresponding quiz questions. Thus, there are a total of 30 quiz questions to assess

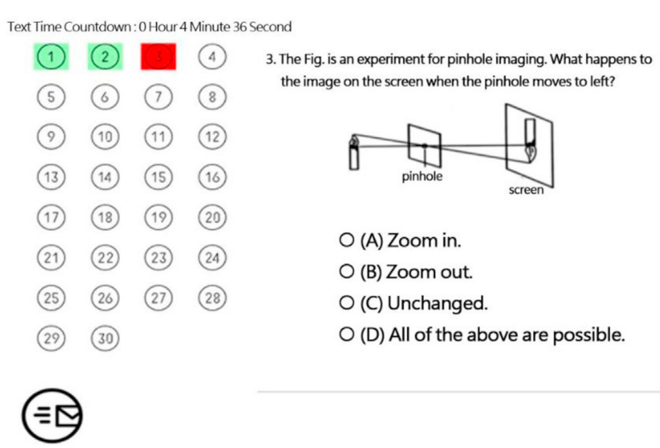

Fig. 1. The test interface of the proposed ITD-DFM.

students' learning performance after the instructor finished the instruction activity.

The main difference between the experimental group and the control group is whether the mechanism of DFM was provided to the students or not. The user interface of the DFM implemented in the ITD-DFM for the experimental group is shown as Fig. 2. After students complete a test, the ITD-DFM will provide an overall assessment report for the experimental group students. The overall assessment report is divided into the left and right blocks. Meanwhile, the left block is further divided into four parts. The first part is the name of the learning unit for the current test, the second part is the correct number of answered questions, the third part is the wrong number of answered questions, and the last part is the review suggestions to the examinee. The block on the right of Fig. 2 is a graphical representation of the ratio of the answer distribution, which is based on the correctness and the response time of the test.

Considering the answer correctness and the response time of a test as the indicators of the designed ITD-DFM was inspired from Verbert et al. (2013) study, summarizing that the tracked data of designing a learning dashboard contain time spent, social interaction, document and tool use, artifacts produced, and exercise results/quizzes. Generally speaking, the test response time will be short if a learner is familiar with the concepts he learned. In contrast, if a learner does not understand the concepts he learned enough, he/she may take a long time to think of the correct answer, or he/she is more likely to make an incorrect answer. Therefore, this study uses the correctness of the answer and the response time on answering the question as the design basis for the test feedback dashboard. Therefore, based on the correctness of each test question in the learning unit, and the comparison between the individual's response time with the average response time of the whole class, the students' test results were divided into A, B, C, and D types. Type A is identified as "skilled (fast and correct)" which means an examinee performs a test with shorter response time and higher correctness than average, type B is identified as "understood but not skilled (slow but correct)" which means an examinee performs a test with longer response time but higher correctness than average, type C is identified as "total incomprehension (slow and wrong)" which means an examinee performs a test with longer response time and lower correctness than average, and type D is identified as "misunderstand or guess the answer (fast but wrong)" which means an examinee performs a test with shorter response time but lower correctness than average.

3.2. The review mechanism in the proposed ITD-DFM

According to the determined four types of test results, students can quickly understand their strength and weaknesses in a test so that they can know how to perform an effective review process. After finishing a test, the teacher suggested students to review the question from the type C first because of the learning concepts corresponding to the type C performed by the examinee with longer response time and lower correctness than average. That is, they are the weakest learning concepts of the examinee. According to the type with the highest number of student's answer results, the review comments and explanations of the overall test results are given, as shown in the left bottom of Fig. 2. Taking Fig. 2 as an example, the left side information shows that the total number of correct answers is four questions, and the wrong answer is 26 questions. According to the proportion of each type in the overall learning results, the student's answer belongs to type C, so the review comments show that "you have a limited understanding of the unit. We recommend you to review the concept of not understanding first, and then strengthen the concept that is not sufficient understanding. If you have time, you can review the already well-understood concepts to deepen your impression." The right side part in Fig. 2 is the ratio of each type of answer, in terms of the size of the circle area. When the mouse cursor is moved to each circle, the wrong question numbers for that type will be displayed. The student can select any type of questions to review first. For example, if the student wants to review from the part of total incomprehension

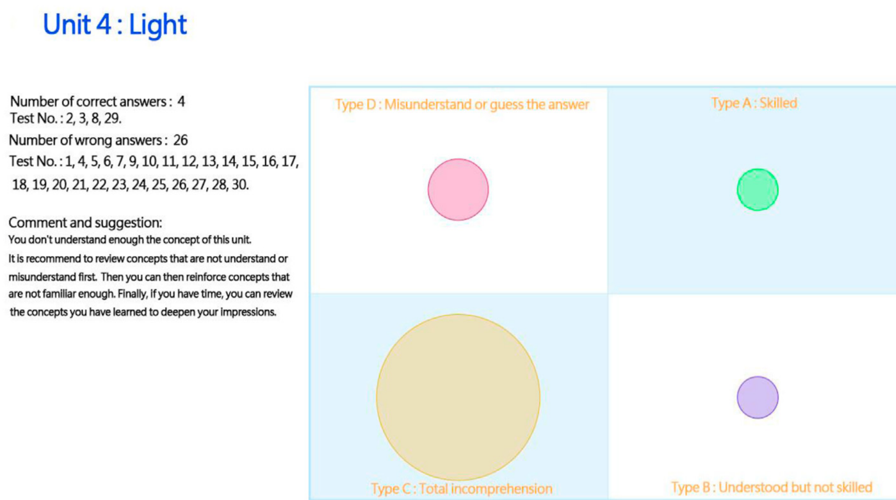

Fig. 2. The user interface of the DFM implemented in the ITD-DFM for the experimental group.

(type C), he can click on the orange circle and the system will change to the review interface, as shown in Fig. 3.

Besides, if the student's answer belongs to type A, the comment will be displayed as "according to the results of your answer, your overall learning outcomes in this unit is great, and you can strengthen the other concepts that are not yet fully understood. Finally, you can review the already well-understood concepts to deepen your impression." If the student's answer belongs to type B, the comment will show up as "you understand the concept of this unit, but it is not enough, you need to strengthen it. We suggest that you first review the concepts that you do not fully understand, and review the concepts that are not sufficient understanding. Then strengthen the already understood but not very skilled parts, and if you have time, you can review the already well-understood concepts to deepen your impression." Otherwise, if the student's answer is type D, the comments will show up as "Although your answer is quick, the answer is wrong. It may be a misunderstanding or a random guess. We recommend you to review the concept of not understanding first, and then strengthen the already understood but not very skilled parts. If you have time, you can review the already well-understood concepts to deepen your impression."

Overall, the interactive features designed in the developed ITD-DFM include that the ITD-DFM provides an interactive test dashboard simultaneously with the interactive mechanisms of diagnosing students' test answer types and providing comments and explanations of the overall test results to guide each learner's learning reflection and review based on simultaneously considering the indicators of test response time and correctness to assist students' learning.

4. Research methodology

This section explains the adopted research methodology of how to examine the differences in learning performance, physics self-efficacy,

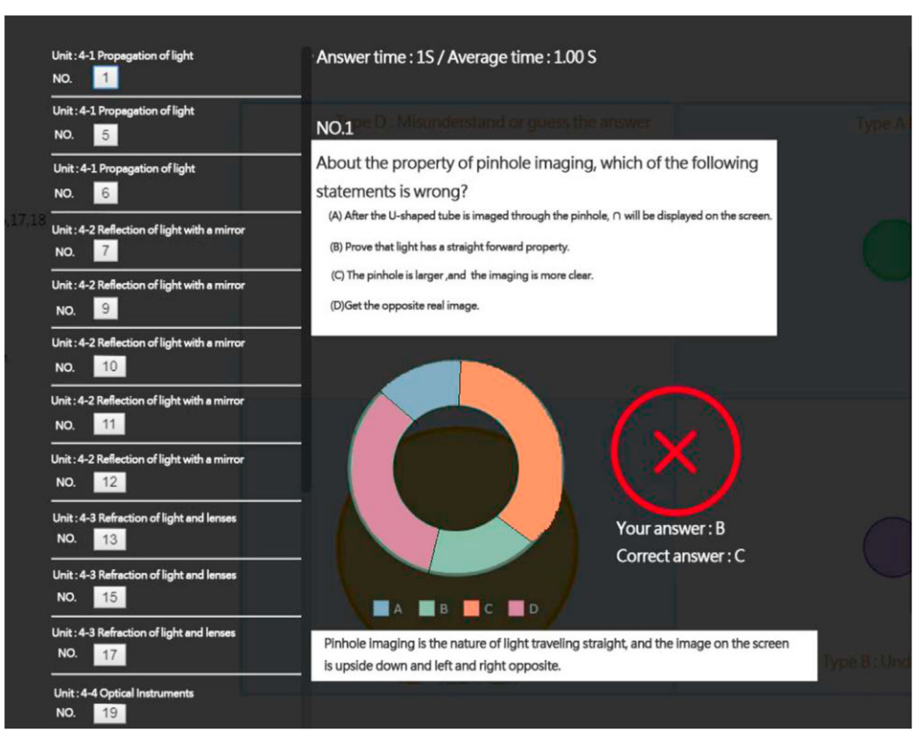

Fig. 3. The review interface of the proposed ITD-DFM.

and technology acceptance between the research participants who respectively used the developed ITD-DFM and the TTS-NDFM to support physics course's teaching and learning for the learning unit of light.

4.1. *Research participants*

This work applied the quasi-experiment nonequivalent control group design to examine the research questions of this study because randomly recruiting research participants to take part in an instruction experiment

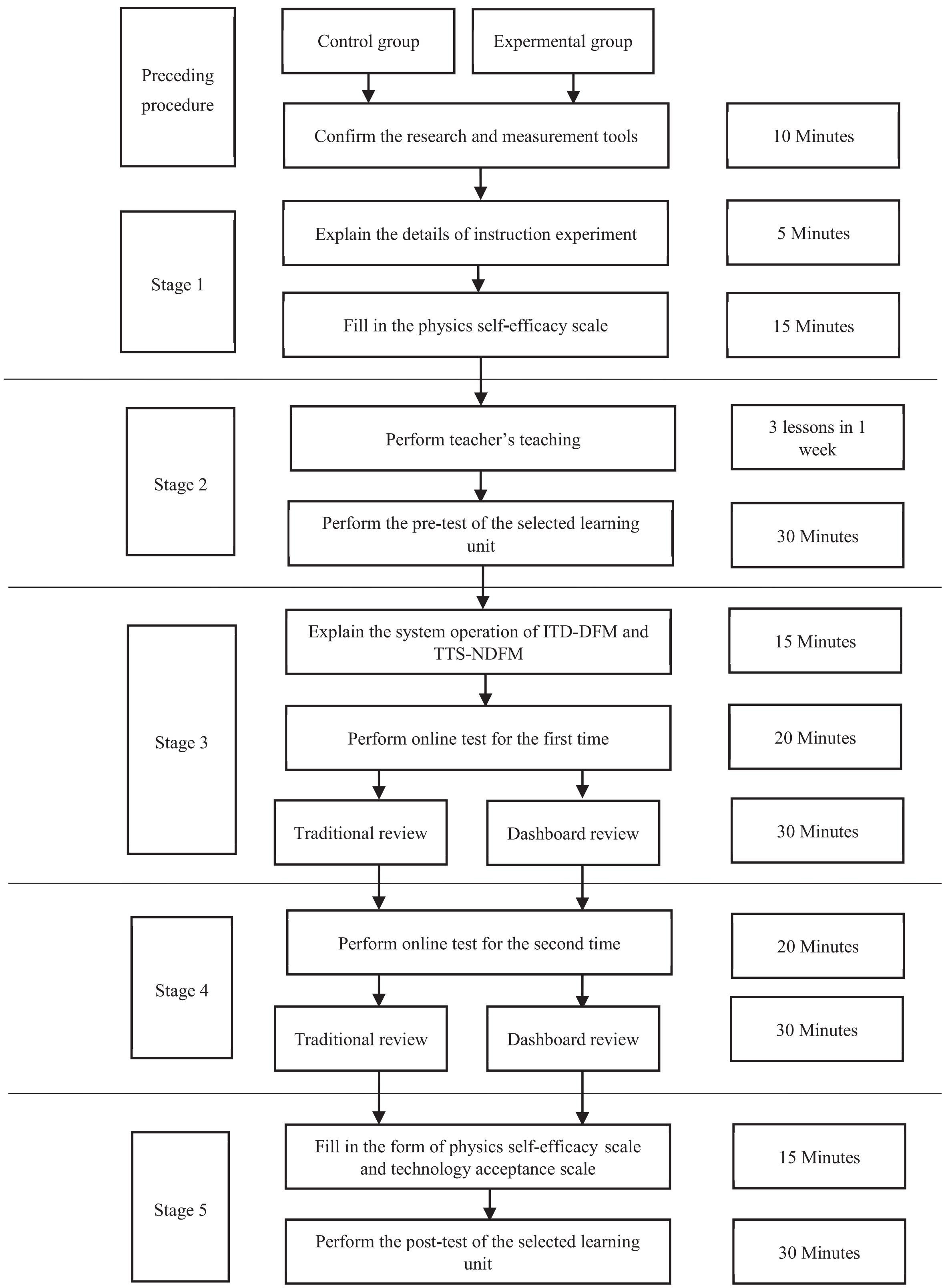

Fig. 4. The learning procedures of the designed instructional experiment.

is a difficult task in actual teaching scenarios. A total of 50 Grade 8 students from two classes of a junior high school in Taiwan were recruited as the research participants. A total of 26 students (15 males, 11 females) from one class were randomly assigned to the experimental group using the proposed ITD-DFM to support physics course's teaching and learning for the learning unit of light, while a total of 24 students (12 males, 12 females) from the other class were assigned to the control group using the TTS-NDFM.

To examine whether there were significant differences in the learning performance, physics self-efficacy, and technology acceptance between the learners with different prior knowledge levels of both groups that use different test systems, the learners of both groups whose pre-test scores in the learning unit of light are higher and lower than the average score of the learners in both groups were respectively identified as high and low prior knowledge levels' learners. In total, the experimental group respectively involves 14 and 12 students with high and low prior knowledge levels, whereas the control group respectively involves 18 and 6 students with high and low prior knowledge levels.

4.2. Experimental procedure

The instruction experiment lasted for 2 weeks. In the first week, all the students in both the groups were taught the learning unit of light by the same teacher in a physical classroom. In the second week, two rounds of test and review activities were conducted in a computer classroom. Fig. 4 shows the detailed learning procedures in the instruction experiment. The experimental procedures can be divided into the preceding procedure and five stages. In the preceding procedure, the used research and measurement tools were confirmed with 10 min, including the learning performance test sheets, online test questions, physics self-efficacy scale, and technology acceptance scale. In the first stage, both the groups were given 5 min to explain the details of the entire instruction experiment. After that, both the groups were given a total of 15 min to fill in the physics self-efficacy scale. In the second stage, all the participants in both the groups learned the same learning unit with the traditional lecture teaching method and were taught by the same teacher. The teacher's teaching was arranged in three lessons of a week. Then, the learners took 30 min to finish the pretest of the selected learning unit. At the beginning of the third stage, the researcher first explained the system operation of the ITD-DFM and TTS-NDFM for the learners in both the groups with a total of 15 min. Then, the students of both the groups took the first time of online tests for a total of 20 min. After the test, the students of both the groups had a total of 30-min review time. The experimental group used the ITD-DFM for review, while the control group used the TTS-NDFM for review. The fourth stage is the second time of the online test. The test time and review time of the fourth stage is the same as the third stage, which is a total of 20 min and 30 min, respectively. After the end of the fourth stage of the instruction experiment, the fifth stage was immediately proceeded. The participants were invited to fill in the form of a physics self-efficacy scale and technology acceptance scale with a total of 15 min. After that, a post-test for the learning unit of light was taken with 30 min. Besides, this study also conducted a semi-structured interview with several learners randomly selected from the experimental group.

4.3. Research tools

4.3.1. The test system used for the control group

The ITD-DFM used for the experimental group has been explained detailedly in section 3. This section aims to explain the TTS-NDFM used for the control group. The TTS-NDFM is a traditional test system that only provides students' test results and does not conduct an analysis of identifying learners' test types based on the response time and correctness of the students' answers. The user interface of the TTS-NDFM is divided into three blocks and is shown as Fig. 5. The first block at the top of the figure shows the name of the learning unit for the current test. The second block

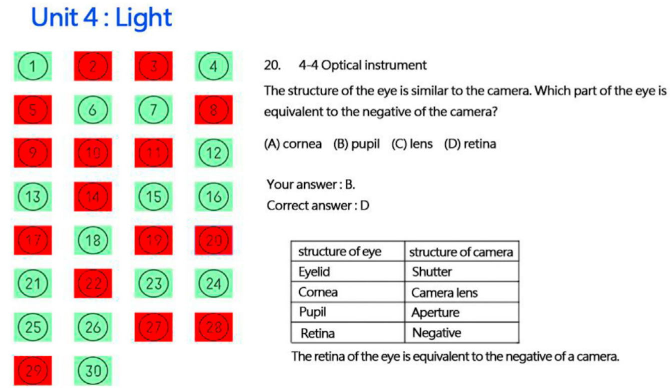

Fig. 5. The test system named as TTS-NDFM used in the control group.

at the left of the figure is the test question number and marks the correct answers with green color and the incorrect answers by red color. The third block at the right of the figure is the topic area. Corresponding to the selected question number, the third block presents the learning concept, question, answer options, correct answer, and detailed explanations for the test question. Students can choose the topic to view according to the question number, but they cannot know the overall test analysis results and the familiarity level of each topic and concept.

4.3.2. Learning performance test sheet

The pretest and posttest sheets that respectively contain 30 questions selected from the test bank developed by National Academy for Educational Research in Taiwan according to five subunits of light including light propagation, light reflection and mirror, light refraction and lens, optical instrument, and light and color were used to assess the research participants' prior knowledge and learning performance. To let the pretest and posttest sheets contain uniform question distribution in the five subunits, each subunit was involved six questions. The evaluation method is scoring one point for each correct answer and a maximum of 30 points. The questions of the pretest sheet are the same with the posttest sheet, but the order of the questions and each question's answer sets were rearranged to avoid memory effect during the test process.

4.3.3. Physics self-efficacy scale

The physics self-efficacy scale used in this study mainly refers to the physics self-efficacy scale (PSES) designed by Çaliskan, Sezgin, and Erol (2007) that is a self-administered measure to assess physics self-efficacy beliefs regarding to one's ability to successfully perform physics tasks in a physics classroom. The scale is divided into five parts: (1) physics self-efficacy towards solving physics problems with a total of 10 questions, such as "I can analyze the context of a physical problem," (2) physics self-efficacy towards physics laboratory with a total of 7 questions, such as "My learning performance in a physical course is as good as that in other courses," (3) physics self-efficacy towards physics achievement with a total of 4 questions, such as "When necessary, I can recall the basic knowledge learned from a physical course," (4) physics self-efficacy towards the application of physics knowledge with a total of 6 questions, such as "I can ask questions in a physics course," and (5) the self-efficacy of memorizing physics knowledge with a total of 3 questions, such as "I can understand the experimental procedure recorded in an experiment manual." The scale is scored using the 5-point Likert option, including "strongly disagree," "disagree," "normal," "agree" and "strongly agree." A higher score indicates that learner has better physics self-efficacy. The Cronbach's α value of the overall scale was as high as 0.94, indicating that the scale had good reliability and was suitable for use as a research tool.

4.3.4. Technology acceptance scale

To elicit the research participants' opinions about using the ITD-DFM and the TTS-NDFM to support learning for the unit of light in junior high school's physics course, they were invited to fill in a technology acceptance questionnaire after the instruction experiment. The analysis of their responses revealed the research participants' subjective perceptions of the extent to which the ITD-DFM and the TTS-NDFM improved their reading performance and the perceived difficulty of operating the two systems. The technology acceptance questionnaire of Hwang, Yang, and Wang (2013) was used and some sentences modified to conform to the present research, and a Likert six-point scale was used. The questionnaire with two dimensions includes 6 questions that examine the perceived usefulness of the system and 13 questions that examine the perceived ease of use of the system. The Cronbach's α values of the subscales of the perceived usefulness of the system and the perceived ease of use of the system were 0.95 and 0.94, respectively, both indicating good reliability.

5. Experimental results

5.1. Analysis of the differences in learning performance for both groups

One-way analysis of covariance (ANCOVA) was conducted to examine whether there was a significant difference in the learning performance between the two groups. Before conducting ANCOVA, the homogeneity of regression coefficients was analyzed first, using the pretest scores as covariates in the analyses. The results confirmed the assumptions of homogeneity of regression coefficients for the posttest ($F = 0.055$, $p = 0.815 > 0.05$). One-way ANCOVA was then proceeded and the results are shown in Table 1. After controlling the effect of learners' pretest scores, the experimental group performed significantly better in the posttest ($F = 5.32$, $p = 0.026 < 0.05$, $\eta^2 = 0.10$) than that of the control group. The results show that the ITD-DFM could better promote the learning performance than the TTS-NDFM. The calculation result of the effect eta squared (η^2) is 0.10, according to Cohen's study (1988). An eta squared (η^2) ranged between 0.06 and 0.14 indicates that the difference in learning performance between both the groups is moderately different.

5.2. Analysis of the difference in physics self-efficacy for both groups

Table 2 shows the results of the descriptive statistics of the physics self-efficacy for both groups. ANCOVA was conducted to examine whether there was a significant difference in the physics self-efficacy between the two groups. Before conducting ANCOVA, the homogeneity of regression coefficients was analyzed first, using the pretest scores as covariates in the analyses. However, this study found that the assumption of homogeneity of regression coefficients is violated ($F = 29.52$, $p = 0.000 < 0.05$). That is, ANCOVA could not be proceeded. This study thus used the post-test score minus the pre-test score as the gain score to compare the differences in the physics self-efficacy improvement between both groups based on the independent samples *t*-test. The results of the analysis are shown in Table 3. The results show that the difference in the gain score between the two groups was significant ($t = 10.40$, $p = 0.000 < 0.05$, $dppc2 = 2.97$), and the experimental group was significantly superior to the control group. That is, the use of the ITD-DFM could promote students' physics self-efficacy. The effect size d_{ppc2} (Morris, 2008) is 2.97. Cohen (1988) suggested that d = 0.2 can be considered with a small effect size, 0.5 represents a medium effect size, and 0.8 represents a large effect size. Therefore, the physics self-efficacy improvement of the experimental group and the control group is largely different. That is, the physics self-efficacy of the learners in the experimental group was significantly enhanced due to using the ITD-DFM to support their learning in a physics course.

Table 2
The descriptive statistics results of the physics self-efficacy for both groups.

Physics self-efficacy	Group			
	Experimental group (n = 26)		Control group (n = 24)	
	Mean	Standard deviation	Mean	Standard deviation
Pre-test	2.24	0.57	2.50	0.38
Post-test	3.98	0.54	2.77	0.42

5.3. Analysis of the difference in technical acceptance for both groups

The difference in technical acceptance between both groups was assessed. Table 4 shows the analytical results based on the independent samples *t*-test with the questionnaire score of technical acceptance filled out by the learners of both groups. The results show that the difference in technical acceptance between both groups differed significantly ($t = 3.30$, $p = 0.002 < 0.05$, $dppc2 = 0.94$). According to Cohen (1988), Cohen's *d* effect size that is higher than 0.8 can be regarded as "large." This indicates that the technical acceptance of the experimental group using the proposed ITD-DFM to support learning in a physics course had significantly better than that of the control group using the TTS-NDFM.

5.4. Analysis of the difference in learning performance, physics self-efficacy, and technology acceptance between learners with different prior knowledge levels in both groups

5.4.1. Analysis of the difference in learning performance between learners with different prior knowledge levels in both groups

One-way analysis of covariance (ANCOVA) was conducted to examine whether there was a significant difference in the learning performance between the learners with different prior knowledge levels in both groups. Before conducting ANCOVA, the homogeneity of regression coefficients was analyzed first, using the pretest scores as covariates in the analyses. The results confirmed the assumptions of homogeneity of regression coefficients for the posttests of the learners with high prior knowledge level ($F = 0.04$, $p = 0.849 > 0.05$) and with low prior knowledge level ($F = 1.03$, $p = 0.337 > 0.05$) in both groups. One-way ANCOVA was then proceeded and the results are respectively shown in Tables 5 and 6. After controlling the effect of learners' pretest scores, neither the learners with high prior knowledge level or low prior knowledge level in the experimental group did not perform significantly better in the posttest ($F = 2.36$, $p = 0.135 > 0.05$, $\eta^2 = 0.08$; $F = 0.82$, $p = 0.378 > 0.05$, $\eta^2 = 0.05$) than those in the control group.

5.4.2. Analysis of the difference in physics self-efficacy between learners with different prior knowledge levels in both groups

For the learners with a high prior knowledge level between both groups, whether the physics self-efficacy differed significantly was

Table 1
The ANCOVA results of learning performance of both groups.

	Learning group				*F*	*p*	*η2*
	Experimental group (n = 26)		Control group (n = 24)				
	Mean	Standard deviation	Mean	Standard deviation			
Pretest	57.88	19.24	72.71	21.47	5.32*	0.026	0.10
Posttest[a]	82.46[a]	2.82	72.75[a]	2.94			

[a] the adjusted posttest scores are used, *p < 0.05.

Table 3
The independent samples *t*-test results of the physics self-efficacy by the gain score of both groups.

	Group				t	p	d_{ppc2}
	Experimental group (n = 26)		Control group (n = 24)				
	Mean	Standard deviation	Mean	Standard deviation			
Gain score	1.75	0.27	0.27	0.65	10.40***	0.000	2.97

***$p < 0.001$.

Table 4
The independent samples *t*-test results of the technology acceptance of both groups.

	Group				t	p	d_{ppc2}
	Experimental group (n = 26)		Control group (n = 24)				
	Mean	Standard deviation	Mean	Standard deviation			
Technology acceptance	3.21	0.46	2.81	0.39	3.30**	0.002	0.94

**$p < 0.01$.

Table 5
The ANCOVA results of learning performance of the learners with high prior knowledge level in both groups.

	Learning group				F	p	$\eta2$
	Experimental group (n = 14)		Control group (n = 18)				
	Mean	Standard deviation	Mean	Standard deviation			
Pretest	72.14	10.32	83.33	10.84	2.36	0.135	0.08
Posttest[a]	93.40[a]	3.05	86.79[a]	2.65			

[a] The adjusted posttest scores are used.

Table 6
The ANCOVA results of learning performance of the learners with low prior knowledge level in both groups.

	Learning group				F	p	$\eta2$
	Experimental group (n = 12)		Control group (n = 6)				
	Mean	Standard deviation	Mean	Standard deviation			
Pretest	41.25	14.45	40.83	9.70	0.82	0.378	0.05
Posttest[a]	59.44[a]	5.30	51.10[a]	7.50			

[a] The adjusted posttest scores are used.

analyzed by ANCOVA due to the assumption of homogeneity of regression coefficients was not violated ($F = 2.05$, $p = 0.174$ >0.05). Table 7 shows the results. The results show that the physics self-efficacy differed significantly between both groups for the high prior knowledge learners (F = 84.73, $p = 0.000 < 0.05$, $\eta^2 = 0.75$). It means that the physics self-efficacy of high prior knowledge learners was improved by the proposed ITD-DFM.

For the learners with low prior knowledge level between both groups, the assumption of homogeneity of regression coefficients is violated ($F = 12.68$, p = 0.006 <0.05). That is, ANCOVA could not be proceeded. Therefore, this study used the independent samples *t*-test to examine the difference in the gain scores between the low prior knowledge learners of both groups. Table 8 shows the results. The results show that the physics self-efficacy of low prior knowledge learners in both groups differed significantly ($t = 7.13$, $p = 0.001$ <0.05, $d_{ppc2} = 3.16$). In other words, the learners with low prior knowledge level in the experimental group had better physics self-efficacy than those in the control group. This work confirmed that the physics self-efficacy of the learners with low prior knowledge level was significantly improved by the proposed ITD-DFM.

5.4.3. Analysis of the difference in technology acceptance between learners with different prior knowledge levels in both groups

Herein, the independent samples *t*-test was used to examine the difference in the technology acceptance between either the learners with high prior knowledge level or low prior knowledge level of both groups. Tables 9 and 10 show the results, respectively. The results show that the technology acceptance of the learners with a high prior knowledge level in the experimental group was significantly higher than that of the

Table 7
The ANCOVA results of physics self-efficacy of the learners with high prior knowledge level in both groups.

	Learning group				F	p	$\eta2$
	Experimental group (n = 14)		Control group (n = 18)				
	Mean	Standard Deviation	Mean	Standard deviation			
Pretest	2.50	0.58	1.92	0.37	84.73***	0.000	0.75
Posttest[a]	4.21[a]	0.12	2.73[a]	0.10			

[a] the adjusted posttest scores are used, ***p < 0.001.

Table 8
The independent samples *t*-test results of the physics self-efficacy of the learners with low prior knowledge level in both groups.

	Group				t	p	d_{ppc2}
	Experimental group (n = 12)		Control group (n = 6)				
	Mean	Standard deviation	Mean	Standard deviation			
Gain score	1.80	0.25	0.49	0.53	7.13**	0.001	3.16

**$p < 0.01$.

Table 9
The independent samples *t*-test results of the technology acceptance of the learners with high prior knowledge level in both groups.

	Group				t	p	d_{ppc2}
	Experimental group (n = 14)		Control group (n = 18)				
	Mean	Standard deviation	Mean	Standard deviation			
Technology acceptance	3.26	0.51	2.83	0.35	2.73*	0.010	0.98

*$p < 0.05$.

Table 10
The independent samples *t*-test results of the technology acceptance of the learners with low prior knowledge level in both groups.

	Group				t	p	d_{ppc2}
	Experimental group (n = 12)		Control group (n = 6)				
	Mean	Standard deviation	Mean	Standard deviation			
Technology acceptance	3.13	0.40	2.70	0.48	2.01	0.061	0.97

learners with a high prior knowledge level in the control group ($t = 2.73$, $p = 0.01 < 0.05$, $d_{ppc2} = 0.98$). However, the technology acceptance of the learners with low prior knowledge level in both groups did not differ significantly ($t = 2.01$, $p = 0.061 > 05$, $d_{ppc2} = 0.97$).

6. Discussion

In this work, an ITD-DFM was developed to facilitate individual learner's learning performance by generating visualized, rich, and high-quality test review information based on simultaneously considering the test response time and correctness of a test. Analytical results revealed that using the proposed ITD-DFM to support learning can significantly promote the learning performance of learners in comparison with TTS-NDFM. Many previous studies (Arnold & Pistilli, 2012; Vieiraa, Parsons, & Byrd, 2018) have indicated that using the digital learning dashboard can help learners improve their learning performance. Particularly, Kokoç and Arif (2019) indicated that learning dashboards could be applied in online courses as teaching assistants to improve the performance of learners in e-learning environments. The ITD-DFM designed using in this study can help students judge their familiarity with the questions according to the response time and accuracy of the students' answers in a test. In other words, students can use the proposed ITD-DFM to understand their problems in learning and to reinforce the learning concepts they are not familiar with. Besides, there are no significant differences in the learning performance of the students with different prior knowledge levels in both groups respectively using the ITD-DFM and TTS-NDFM to support learning. It means that the ITD-DFM had the same effect on learning performance regardless of learners with high or low prior knowledge level. The result is inconsistent with several previous studies (Chen, Chen et al., 2019; Chen, Wang et al., 2019; Liu et al., 2008), indicating that computer-supported learning systems generally provide more benefits in terms of promoting learning performance for the learners with low prior knowledge level in comparison with the learners with high prior knowledge level. For example, Chen, Chen et al. (2019) employed C4.5 decision tree to develop a collaborative reading annotation system with formative assessment and feedback mechanisms (CRAS-FAFM) based on four considered social network indicators, which could forecast the learners with low reading comprehension and suggest them to interact with the learners who are predicted with high reading comprehension performance and infrequently interact in the digital reading activity to enhance their reading comprehension through interactive discussion. Their study found that compared to the collaborative reading annotation system without formative assessment and feedback mechanisms (CRAS-NFAFM), the CRAS-FAFM provides remarkable benefits in promoting the reading comprehension performance and interactive discussion on the discussion level of comparison, discussion, and analysis, particularly for the learners with low prior knowledge level.

The study also found that the progress of the physics self-efficacy of the students who are either high or low prior knowledge level using the proposed ITD-DFM to support learning is significantly higher than those of using TTS-NDFM. This result shows that the ITD-DFM could help students' review and achieve good learning performance through the mechanism of DFM, thereby enhancing their physics self-efficacy. Many previous studies have found that self-efficacy can predict learning achievement in science for high school students (Britner, 2008; Lau & Roeser, 2002). This study also confirmed that physics self-efficacy may facilitate students' learning performance in a physics course. Besides, compared to the TTS-NDFM, the proposed ITD-DFM provides remarkable benefits in promoting technology acceptance, particularly for the learners with high-prior knowledge level. It means that the learners with high prior knowledge level perceived the usefulness and ease of use of the proposed ITD-DFM much more than the students with low prior knowledge level. Namely, the ITD-DFM is an assisted learning tool towards learners with higher prior knowledge level.

Moreover, based on the feedback from the students in the experimental group who participated in a semi-structured interview, most of the students suggested that the proposed ITD-DFM should provide more questions for the learning unit and highlight the key parts of the question for the learners to know where the question focuses on. In terms of interface design, student's suggestions included zooming in on fonts and options to increase user convenience. Besides, students also hoped that in the future, the ITD-DFM will be launched to the cloud platform so that they can freely use and review at any time without restriction used in school classrooms.

7. Conclusions and future work

This study presents a novel ITD-DFM that can provide visualized, rich, and high-quality test feedback as review signals to promote students' learning performance, physics self-efficacy, and technology acceptance based on simultaneously considering the features of summative assessment and self-assessment. Importantly, the DFM can identify examinees' test results as A, B, C, and D types so that examinees can know what learning concepts are "skilled (fast and correct)," "understood but not skilled (slow but correct)," "total incomprehension (slow and wrong)," and "misunderstanding or guessing the answer (fast but wrong)," thus guiding them to make more efficient review process. Moreover, the ITD-DFM had the same effect on learning performance and physics self-efficacy regardless of learners with high or low prior knowledge level, but it provided higher technology acceptance to the learners with high prior knowledge level than did the learners with low prior knowledge level. The implication of the proposed ITD-DFM is successfully to expand the features of traditional summative assessment based on test questions and visualized feedback for review as well as bring the summative assessment into new ground.

This study suggests two future research directions. Firstly, the learning unit of light designed in this study is a physical subject. In general, the physical subject requires more mathematical operations and logical understanding. A future study may consider applying the proposed ITD-DFM to other different disciplines, such as language, history, or geography, to expand its application in different educational scenarios. Secondly, the proposed ITD-DFM did not record the detailed operation behaviors of learners using the ITD-DFM to support learning. Future study can use the new-generation learning process recording technology, xAPI (Experience Application Programming Interface), to record students' learning processes and progress (Manso-Vázquez, Caeiro-Rodríguez, & Llamas-Nistal, 2015) as well as discuss how learners' operation behaviors affect their learning performance, physics self-efficacy, and technology acceptance.

Compliance with ethical standards

To consider the research ethics of the designed experiment that involves recording the test behaviors of the research subjects by using programming technologies, written informed consent was obtained from the research subjects following a full explanation of the experiment. The informed consent letter contains the specific nature of the research, including the data that collect from them, are only for the research, their name will never appear on any data collected, we will provide a unique identification number on their data and that information will remain secure such that only the principal investigator of this study will have access to it, and the collected data that is no longer needed will be destroyed. Moreover, all procedures performed in this study involving human participants were following the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Finally, we certify that there is no conflict of interest in this paper.

Availability of data and material

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Funding

Ministry of Science and Technology in Taiwan, ROC [grant number MOST 106-2511-S-004 -006 -MY3].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In S. Buckingham Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International conference on learning analytics and knowledge (LAK '12)* (pp. 267–270). New York: ACM.

Bajzek, D., Brown, W., Lovett, M., & Rule, G. (2007). *Inventing the digital dashboard for learning*. Waynesville, NC: Association for the Advancement of Computing in Education (AACE).

Britner, S. L. (2008). Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes. *Journal of Research in Science Teaching, 45*(8), 955–970.

Çaliskan, S., Sezgin, G. S., & Erol, M. (2007). Development of physics self-efficacy scale. *AIP Conference Proceedings, 899*, 483–484.

Chen, C. M., Chen, L. C., & Horng, W. J. (2019a). A collaborative reading annotation system with formative assessment and feedback mechanisms to promote digital reading performance. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2019.1636091

Chen, C. M., Wang, J. Y., & Lin, M. (2019b). Enhancement of English learning performance by using an attention-based diagnosing and review mechanism in paper-based learning context with digital pen support. *Universal Access in the Information Society, 18*, 141–153.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: NJ Lawrence Erlbaum Associates, Publishers.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.

Earl, L., & Katz, S. (2006). *Rethinking classroom assessment with purpose in mind. Winnipeg, Manitoba: Western Northern Canadian protocol*.

Ellis, A. E. (2003). Personality type and participation in networked learning environments. *Educational Media International, 40*, 101–112.

de Freitas, S., Gibson, D., Alvarez, V., Irving, L., Star, K., Charleer, S., et al. (2017). How to use gamified dashboards and learning analytics for providing immediate student feedback and performance tracking in higher education. In *Proceedings of the 26th International conference on world wide web companion* (pp. 429–434). International World Wide Web Conferences Steering Committee.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education, 69*, 121–130.

Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). License to evaluate: Preparing learning analytics dashboards for educational practice. In *Proceedings of the 8th International conference on learning analytics and knowledge - LAK'18* (pp. 31–40). Sydney, New South Wales, Australia: ACM Press.

Knight, P. T. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education, 27*(3), 275–286.

Kokoç, Mehmet, & Altun, Arif (2019). Effects of learner interaction with learning dashboards on academic performance in an e-learning environment. *Behaviour & Information Technology*, 1–15. https://doi.org/10.1080/0144929X.2019.1680731

Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment, 8*(2), 139–162.

Liu, H. C., Andre, T., & Greenbowe, T. J. (2008). The impact of learner's prior knowledge on their use of chemistry computer simulations: A case study. *Journal of Science Education and Technology, 17*(5), 466–482.

Magill, R. A. (2007). *Motor learning and control: Concepts and applications*. Boston: McGraw-Hill.

Manso-Vázquez, M., Caeiro-Rodríguez, M., & Llamas-Nistal, M. (2015). xAPI-SRL: Uses of an application profile for self-regulated learning based on the analysis of learning strategies. In *Proceedings of the 2015 IEEE frontiers in education conference* (pp. 1–8). Washington, DC, USA: IEEE Computer Society.

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.

Nyland, R. (2018). A review of tools and techniques for data-enabled formative assessment. *Journal of Educational Technology Systems, 46*(4), 505–526.

Park, Y., & Jo, I. H. (2015). Development of the learning analytics dashboard to support students' learning performance. *Journal of Universal Computer Science, 21*(1), 110–133.

Park, Y., & Jo, I. H. (2019). Factors that affect the success of learning analytics dashboards. *Educational Technology Research & Development, 67*(6), 1547–1571.

Riofrío-Luzcando, D., Ramírez, J., Moral, C., de Antonio, A., & Berrocal-Lobo, M. (2019). Visualizing a collective student model for procedural training environments. *Multimedia Tools and Applications, 78*(8), 10983–11010.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., et al. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*(1), 30–41.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist, 57*(10), 1500–1509.

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., et al. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing, 18*(6), 1499–1514.

Vieiraa, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education, 122*, 119–135.

Vie, J. J., Popineau, F., Bruillard, É., & Bourda, Y. (2017). A review of recent advances in adaptive assessment. In A. Peña-Ayala (Ed.), *Learning analytics: Fundaments, applications, and trends. Studies in systems, decision and control* (Vol. 94). Cham: Springer.

Yu, Y. C. (2017). Teaching with a dual-channel classroom feedback system in the digital classroom environment. *IEEE Transactions on Learning Technologies, 10*(3), 391–402.