

# Kit-Build Concept Map with Confidence Tagging in Practical Uses for Assessing the Understanding of Learners

Jaruwat Pailai, Warunya Wunnasri, Yusuke Hayashi, Tsukasa Hirashima  
Graduate School of Engineering  
Hiroshima University  
Hiroshima, Japan

**Abstract**—An answer of a learner can be interpreted as a learning evidence for demonstrating the understanding of the learner, while a confidence on the answer represents the belief of the learner as the degree of understanding. In this paper, we propose Kit-Build concept map with confidence tagging. Kit-Build concept map (KB map in short) is a digital tool for supporting a concept map strategy where learners can create the learning evidence, and the instructor can access the correctness and confidence information of learners. The practical uses were conducted for demonstrating the valuable of correctness and confidence information in the lecture class. The correctness information was visualized in the control classes, while the correctness and confidence information were visualized in the experiment classes. The observed evidence illustrates that the different information was used for selecting and ordering the supplementary content when the system visualized the different information. The normalized learning gains and effect size demonstrate the different learning achievements between control- and experiment- classes. The results suggest that the confidence information of learner affects the instructor behaviors, which is the positive changing behavior for improving the understanding of their learners. The results of questionnaire suggest that the KB map with confidence tagging is an accepted mechanism for representing the learner's understanding and their confidence. The instructors also accepted that the confidence information of learners is valuable information for recognizing the learning situation.

**Keywords**—Kit-Build concept map; confidence tagging; effect of confidence information; behavior changing of instructor

## I. INTRODUCTION

Knowledge is invisible, but it is possible to observe its effect. The knowledge is often defined as a belief, which is true and justified, while the certainty is an essential component of knowledge [1]. More discussion about the concept of knowledge was described by Hunt who mentioned that a knowledge measurement requires a measuring of correctness and sureness. The quality of knowledge can be represented by the certainty of the answer such as a learner who is sure on the correct answer and a learner who is not sure on the incorrect answer. Besides, the confidence is essential to influence real-life behavior, many decision-making, and learning processes [2]-[8]. Several researchers mentioned the confidence in the various situations. For instance, the confidence can encourage a deeper understanding of the material [9], and also increase

reflection and justification of the answers [10]. Consequently, the answer of learners represents their understanding, and the confidence on their answer indicates the degree of understanding. The answers and its confidence are possible to be utilized as the learning evidence of formative assessment for identifying the information of current learning situation, which the instructor can use to design and provide the feedback as the evidence-based feedback. The value of information will be indicated by an instructor when s/he used the information to provide the feedback for improving learners' understanding in their class.

The Kit-Build concept map (KB map in short) is a digital tool for supporting a concept map strategy, which can identify the correctness of learners-build concept map based on the instructor-build concept map automatically [11]. The learning goal is represented in the form of a concept map for indicating the expectation of the instructor, which the instructor-build map is called a goal map. Learners can construct the learner maps as the learning evidence by connecting the provided components of concept map (as we called "Kit") to form each proposition. The kit is the list of concepts and linking words from a decomposing the goal map. The assessing process of the KB map is to identify the correctness information by using the propositional level exact matching for generating the diagnosis results. The correctness information is available in the diagnosis results, which can be divided into individual-diagnosis results for informing the performance of learners individually and group-diagnosis results for informing an overview of the class. These abilities can help the instructor to reduce the gathering and assessing time instantly. The instructor can use the correctness information in various scenarios such as an intra-class feedback and an inter-class feedback [12], [13]. For instance, the group-diagnosis results identified the incorrect propositions that represent the misunderstanding of the class in only one map. The instructor can find the overview of class easily and prepare to provide the feedback shortly. Accordingly, the ability of the KB map suites to support the instructor for implementing the formative assessment in a classroom situation [14].

In this paper, we propose KB map with confidence tagging for eliciting learning evidence of learners and informing the correctness and confidence information to the instructor. The confidence tagging is integrated into the structuring task of the KB map, which learners can construct the map to represent

their understanding and identify their confidence on each unit of meaning. A completed proposition, which is able to tag the confidence, comprises one connected linking word between two concepts. The confidence of an answer is simplified in the form of confidence- and unconfidence-value, which the learner can assign to every complete proposition. Thus, the system can elicit learning evidence that includes the understanding of learners and the degree of the understanding in the gathering process. The confidence information of learners is utilized in the diagnosis results of the KB map for visualizing the degree of learner's understanding. Therefore, we present the practical uses of the KB map with confidence tagging in the classroom situations when the instructors implement the formative assessment in the lecture classes for illustrating the encouragement of correctness and confidence information in their instruction. Five paired classes were conducted in the practical uses, which each paired class was conducted by the same instructor, the same lecture topic, and two different classes. Only the correctness information was provided to the instructors of five control classes as a control group, while both correctness and confidence information were provided to the instructors of five experimental classes as an experiment group.

The investigation procedure focuses on the different behavior of the same instructor when s/he received the different information on the diagnosis results. From this procedure, we assume that the confidence information of learners effects on the supplementary content ordering of the instructor. The actual ordering of supplementary lecture was used as observed evidence to indicate how the instructor used the correctness and confidence information. Moreover, the normalized learning gains of class and the effect size demonstrate the different learning achievement between both groups, which can illustrate that the correctness and confidence based feedback of the experiment group can contribute the improvement of learning achievements better than the correctness based feedback of the control group in several classes. The learners of the experiment group have an ability to discriminate and interpret their understanding between correctness and confidence better than the learners of the control group significantly. Analysis of change of proposition type presents that the unconfident propositions are easier to be changed than the confident proposition. Finally, the questionnaire presents that the KB map with confidence tagging is an accepted mechanism. The learners accepted the mechanism for presenting their understanding as propositions and for tagging their confidence to each proposition. The instructors accepted that the confidence information of learners was the valuable information to identify learning situation and identify the degree of learners' understanding.

This paper is structured as follows: Section II demonstrates the utilizing of correctness and confidence information for classifying an answer of a learner. The formative assessment in a lecture class for improving the learning achievements, and the KB map for assessing the understanding of learners are also described in this section. Section III presents KB map with confidence tagging, the practical uses of the KB map in a lecture class, and the description of procedure. The results section, outlined in Section IV presents the observed evidence of the instructors and the learning achievements of learners.

Section V is the discussion about the effect of confidence information of learners on the instructors' behavior. Lastly, Section VI is the conclusion of this study.

## II. BACKGROUND

### A. An Assessment by using Correctness and Confidence

The confidence was used to ensure the performance of learning outcomes as the quality of knowledge or the actual performance [15] as one of assessment criteria. Confidence based learning promotes a fusion of correctness and confidence to identify the answer of learners in four quadrants. There is a definition of correctness and confidence for referencing following:

- *Correctness* is the justification of an answer, which consists of a correct answer and an incorrect answer.
- *Correct- or incorrect-* answer is justified by the criteria.
- *Confidence* is the certainty of an answer, which can be simplified the values as confidence and unconfidence.
- *Confidence- or unconfidence-* of the answer is stated by learners on their answer.

The two-dimensional assessment process was used to classify the answer into four quadrants based on the correctness and confidence simultaneously. The four quadrants of two-dimensional assessment following:

- A correct answer with confidence.
- A correct answer with unconfidence.
- An incorrect answer with confidence.
- An incorrect answer with unconfidence.

Several researchers have already proposed the scoring method based on the correctness and confidence for promoting the critical awareness and self-assessment [16]-[19], for instance, Certainty-based Marking (CBM), Confidence-based Scoring (CBS), and Certainty-based Assessment (CBA). The correct answer that learner has a confidence can get the score more than the correct answer with unconfidence, while the learner can get some score on the incorrect answer when s/he has no confidence on the answer. Zero scores or penalty score is given to the incorrect answer with confidence. The task to identify the confidence of learners on their answer is provided to learners in various strategies such as the answering of descriptive question, True/False question, or the multiple-choice question. The different values of confidence were applied to the scoring method. For instance, the two different values of sureness consist of sure and not sure, or the three different levels of certainty consist of low, middle, and high.

### B. Formative Assessment

A formative assessment provides an opportunity to improve learning achievements, which is different from evaluation in the form of a summative assessment. The key questions of formative assessment following: "Where are learners going?", "Where are learners now?" and "How to close the gap?" [20]. The information through the assessment can encourage the instructor for giving the feedback to improve the learners'

understanding in a timely manner, which is the most efficient feedback [21]. The interaction based on formative information is the formative assessment key feature [22]. The gathering and assessing the learning evidence for providing the feedback are the processes of completing the formative assessment, and are also creating an opportunity for improving learning achievements concurrently. Thus, the formative assessment approach is used to monitor the learning of learners for providing ongoing feedback, which is a key for helping the learners to achieve a learning goal. The learning goal indicates the answer of “where learners are going?” question and can be used as criteria for examining the learner’s knowledge. Subsequently, the learning evidence is assessed by the criteria to identify the correctness for responding “where are learners now?” question. In other words, the correctness of the evidence can inform the learning gap based on learning goal and the evidence of learners. The remaining requirement is “how to close the gap?” question, which can be solved by feedback. The instructor’s feedback is provided in a lecture class as a group feedback for improving the understanding of learners when the instructor duels with a large number of learners. Moreover, the individual feedback is possible to provide in the proper situation such as a focused class with a small number of learners. For instance, the instructors can give the feedback as the supplementary lecture based on the overview of their class, while the different feedback can be provided to some learners individually according to each learner’s misunderstanding after finished class. Thus, the implementation of formative assessment is a completion of the formative assessment cycle following gathering and assessing learning evidence and providing the feedback. It can create a chance to improve learning achievements in every cycle.

### C. Kit-Build Concept Map

Concept maps are graphical tools that are used to represent and organize knowledge [23]. A proposition is constructed by connecting two concepts via a relation with linking word for representing a unit of meaning. The propositions are a core component of measuring a map score. In education areas, concept maps strategy is utilized to represent and assess knowledge of learners in classes as the learning evidence. An instructor can gain the information of a classroom situation then give the feedback based on the information in various situations such as the using individual or group discussion can contribute self-awareness of learners [24], or the using of concept maps as a formative strategy to find discrepancies based on the criteria map before instructor gives the feedback to learners [25]. The concept map strategy is simple to use, effective, and satisfy on problem-solving in classroom situations [26], [27]. Accordingly, the concept map is an effective strategy in a classroom situation that affects achievements and interests of learners. Although the traditional lecturing contributed learning achievements and meaningful learning in the classroom situation, the concept map can significantly improve learning achievements of learners when compared with the lecturing and is also more effective than the traditional lecturing in encouraging meaningful learning [28]-[30].

The KB map is a digital tool for supporting a concept map strategy, which includes a construction tool where users can

construct concept maps and an automatic concept map assessment where the system can report diagnosis results [11]. The different type of concept map is available in different tasks and different meaning, which are connected to each other to form a reasonable relationship respectively. The primary map is a goal map (Fig. 1) that an instructor builds a traditional concept map as criteria for indicating a learning goal of the class. The learning evidence is constructed by the learners from connecting provided components. The provided components are the decomposed components of the goal map as a “kit” (Fig. 2) in the form of a list of concepts and linking words. A learner can connect a linking word between two concepts to form a proposition as a unit of meaning, and then all propositions become a learner map (Fig. 3) for representing their understanding as an answer. Subsequently, the correctness of each learner map is indicated by the propositional level exact matching automatically. The correctness information is reported in the form of an individual-overlay map, an individual-difference map, and a similarity score. The individual-overlay map contains the similarity score and modified visualization of learner map where the displaying of the line connection of the correct proposition is different from the incorrect proposition. The individual-difference map is a visualizing only the mistake of learners in the form of three types of error link, and the linking word of correct propositions are disappeared in this map.

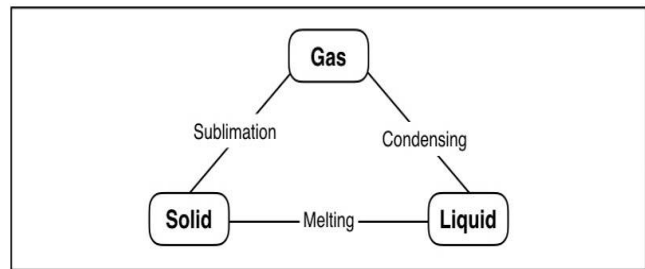


Fig. 1. An example of a goal map.

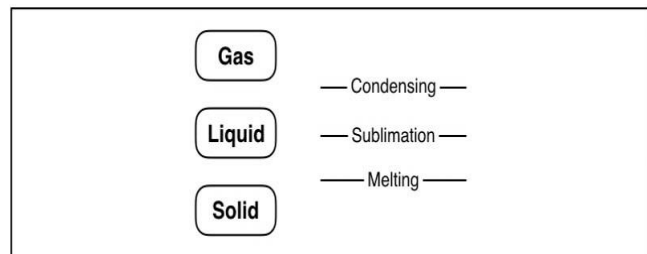


Fig. 2. An example of a kit.

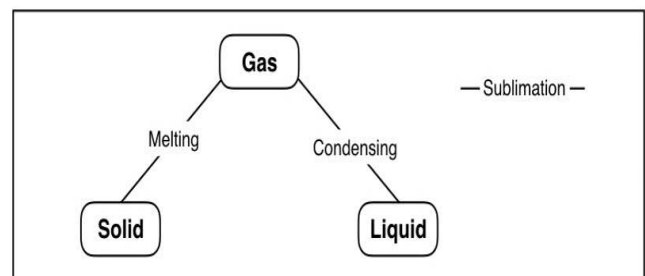


Fig. 3. An example of a learner map.

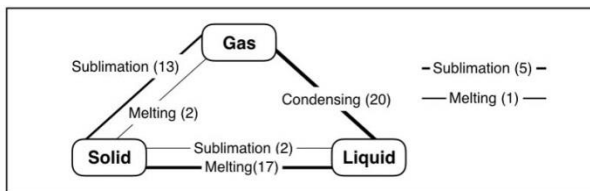


Fig. 4. An example of a group map.

The three types of error link consist of excessive links, leaving links, and lacking links. The link that is used to connect two concepts in learner map but at least one concept which is different from the goal map is called excessive link. The link that is not connected to any concept is called leaving link. The lacking links are used to call the link that is in the goal map but does not exist in the learner map, which the lacking link is the correcting of the excessive link or the leaving link. These are the individual-diagnosis results of the KB map.

Moreover, the advantage of the KB map is group-diagnosis results [31], [32]. A significant component of the KB map is the “kit” which is provided to all learners for constructing learner maps. Thus, all of the learner maps are constructed based on the same components, and overlaying all of the learner maps can be formed as the group-diagnosis results. The group map (Fig. 4) presents the overview of learners’ understanding by visualizing the difference of line weight and tagged the number of learners according to the constructors of each proposition.

In the group-goal difference map, the concepts will be located as same as the concepts in the goal map and only relations of mismatch propositions are visualized. The group-difference map visualizes three types of error link as same as the individual-goal difference map. The excessive link is represented in the form of the solid line which the link is connected with two concepts. It can identify the relations that learners confused or misunderstood, and the tagged number presents the number of learners who constructed the link. The leaving link is also represented in the form of the solid line which the link is not connected with any concept, and indicates that the learners do not understand the linking word. The tagged number means the number of learners who do not use the link to connect with any concept. The dashed line represents the lacking link which is an error correction for displaying the correcting information of excessive- and leaving- links. The tagged number of lacking link is the total number of excessive link and leaving link, which related to the weight of line. The more tagged number in each relation will represent with a thicker line. For instance, an example of a group-goal difference map is shown in Fig. 5. “Melting (3)” dashed line is the lacking link while “Melting (2)” solid line is the excessive link, and “Melting (1)” solid line is the leaving link.

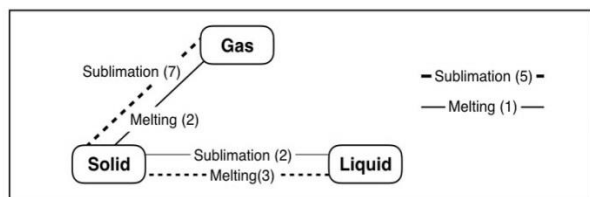


Fig. 5. An example of a group-goal difference map.

In addition, a filtering function of the Kit-Build analyzer can provide more efficient investigation by adjusting the intensity of three types of error link. The filtering function of group-diagnosis results is more explicit with the line weight, which a filtering tool can help the instructor filter out a few error links and keep the big number of error links. A thickness line and a number in parenthesis refer to the number of learners who connect those links. The link of each proposition is available for clicking to discover the learners who are the constructor of the link. Hence, the learner maps will be evaluated through the propositional level exact matching methodology that is the procedure for reporting individual-diagnosis results. The system can provide the additional procedure for reporting the group-diagnosis results at the same time.

Providing the components of the concept map is a kind of “closed-end” approach which is a realizing the automatic diagnosis of the concept map built by a learner [33]. The learner maps of the KB map are composed of the same components with the goal map. Thus, it is possible to detect the difference between them in the form of the diagnosis results. The learners are able to make a map in the limitation of providing parts, which is different from the traditional concept maps where learners can create concept map components by themselves. Therefore, the learners deal with only recall and understanding level in Bloom’s taxonomy [34]. The KB map can utilize in the aspect of confirming the understanding between the instructor and learners in classroom situations with the benefit of the automatic assessment for implementing formative assessment. In addition, the related studies presented the contribution of the KB map on learning effect [35]-[38]. The contribution of the KB map framework has been researched in reading comprehension topic where a direct interaction between the digital tool and learners has been examined. The results show that the KB map can help the learners to retain and recall the information for the longer period of time. The provided components illustrate the effective towards memory as same as the traditional concept map when the learning materials have the clear structure. The arrangement of the KB map on formative assessment also illustrates that an instructor used the suggestion of the diagnosis results for improving learning achievements [12]-[14].

For identifying the degree of learner’s understanding, the confidence identification of learners is the necessary task to indicate their confidence on each unit of meaning. The confidence tagging is utilized to facilitate the gathering of confidence information. The learners can indicate the degree of their understanding on the learner maps, and the diagnosis results also can inform the degree of the learners’ understanding to the instructor. Thus, the KB map with confidence tagging was developed to gather and assess the learning evidence for visualizing both learner’s understanding and the degree of learners’ understanding.

### III. METHODOLOGY

#### A. Kit-Build Concept Map with Confidence Tagging

For gathering learning evidence and identifying the degree of learner’s understanding, the KB map with confidence

tagging was developed for eliciting learning evidence, and associating the correctness and confidence information. In this study, the KB map is reinforced by uniting with the confidence tagging, which is a mechanism for representing learner's understanding on lecture content, and identifying learner's confidence on each proposition of a learner map. The confidence tagging is integrated into the structuring task where the learner constructs a learner map, and a tagging tool (Fig. 6) appears when two concepts and a linking word are connected as a completed proposition. Learners are required to identify their confidence by selecting "sure" or "not sure" on each completed proposition. It is also expected that the tagging task promotes learners to reconsider about their proposition again. The confidence values include "sure" for stating the certainty on the proposition, and "not sure" for indicating unconfidence on the proposition and the system allows the learners to change the values freely. If the learners disconnected the link of the completed proposition, the confidence tagging tool of the link would be disappeared, and the confidence value is reset then. The learners have to identify the confidence value again even they constructed the same proposition after disconnecting. Accordingly, the structuring task of learners can gather the answer of learners and confidence on their answer. Through this task, the system is able to gather the correctness and confidence information of each proposition in all learner maps, and then, the results of the diagnosis about the correctness and confidence are visualized at the same time.

Fig. 7 shows an example of individual-overlay map and Fig. 8 shows an example of a group-difference map, where the correctness and confidence information are reported to the instructor. An additional visualization is a confidence badge. The badge is added into the linking word to indicate the confidence of learners on the link. For instance, a dark tone badge on the dashed line illustrates the incorrect answer with confidence in the individual-overlay map (Fig. 7) of individual-diagnosis results, while a light tone badge on the solid line represents the correct answer with unconfidence.

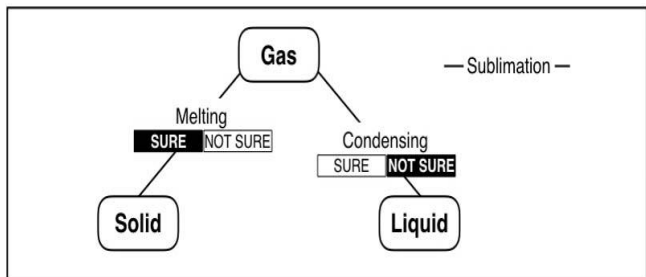


Fig. 6. An example of a learner map with confidence tagging.

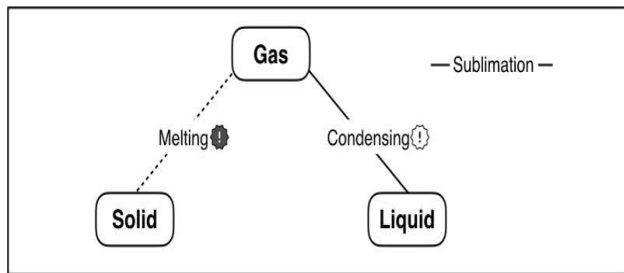


Fig. 7. An example of an individual-goal overlay map.

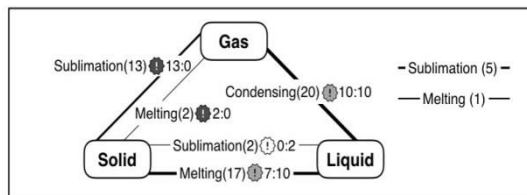


Fig. 8. An example of a group map with confidence information.

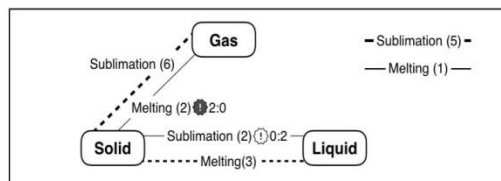


Fig. 9. An example of a group-goal difference map with confidence information.

On the other hand, the mismatch propositions are visualized in group-goal difference map (Fig. 9) of group-diagnosis results where the excessive link indicates the incorrect answer and the lacking link represents the correcting information. A dark tone badge on the solid line illustrates the excessive link with confidence, while a light tone badge on the solid line represents the excessive link with unconfidence. The group-diagnosis results has more details about the confidence information, which the color tone of the badge is varied according to the number of learners who have confidence against unconfidence on the same proposition. For instance, the darkest tone badge has appeared on the link that all of the constructors pressed on "sure" value. A middle tone badge has appeared on the link that the number of "sure" and "not sure" values are equal. The lightest tone badge appeared on the link that no one "sure" on the link. Another indicator is a tagged number of confidence information on the right-hand side of the badge. The colon is punctuation mark for separating the number of learners. The number of learners who pressed on "sure" is displayed on the left-hand side of the mark, while the right-hand side number displays the number of learners who press on "not sure." Fig. 9 shows an example of a group-difference map, where the correctness and confidence information are visualized.

### B. Practical Uses of Kit-Build Concept Map in Lecture Class

The practical uses of the KB map with confidence tagging are an implementation of formative assessment in lecture class for investigating the encouragement of the correctness and confidence information. The instructors can recognize the current learning situation for selecting and ordering the content of supplementary lecture through the analyzer of the KB map with confidence tagging. The participants are three instructors from three different schools, and learners from three different elementary schools who study in the fourth-, fifth-, and sixth-grade. The instructor of fourth grade conducted one practical use, the instructor of fifth grade conducted two practical uses, and the instructor of sixth grade also conducted two practical uses. Ten basic science classes of five paired class are separated into five control classes and five experiment classes. The arrangement of the KB map on formative assessment was used in the practical uses of this study following [14]: the first step is the general scenario of the lecture class, the instructors

created lecture contents and then constructed a goal map for indicating a learning goal of the class. The next step is to give the lecture to learners in a class period. During the lecture, the instructor checks the learner's understanding by requesting learners to construct learner maps and identify their confidence. Then, the diagnosis results are provided to the instructor immediately for informing about current understanding of learners. These steps are gathering and assessing the evidence of learners. The fifth step is to provide intra-class feedback during the class period, which requires an instant practical information for capturing an overall understanding of class. This requirement is responded by the group-diagnosis results that include the group map which can inform the common understanding, and the group-goal difference map which can inform the common misunderstanding of class in one map. Even the inter-class feedback of the sixth step was ignored in the practical uses of this study; we have an additional short discussion session with the instructors after finished classes for summarizing the classroom situation. Fig. 10 illustrates the arrangement of the KB map on formative assessment in a classroom situation.

The supplementary lecture is a feedback of the instructors in the lecture class, which a supplementary content should correspond with the misunderstanding of learners. Even the diagnosis results can identify the understanding and the misunderstanding of learners, the instructor still remains to be the most influential of the class who select the content of the supplementary lecture to raise the understanding of learners as a fulfilling the gaps. The valuable of correctness and confidence information investigation focusses on the behavior of instructors in selecting and ordering the supplementary lecture when the instructor received the different the diagnosis results. The correctness information is also available in the control group, while both the correctness and confidence information are available only in the experiment group. The excessive links of the group-goal difference map present the correctness information, indicate an overview of the incorrect answers, and represent the misunderstanding of learners. The number of excessive links was generally used to order the content of the supplementary lecture. The location of each excessive link was also used for ordering the excessive links that have an equal amount of the constructors (unordering of correctness information). Hence, an assumption of the control group is that the instructor selects the excessive links to provide the supplementary lecture following the correctness information and the location of visualization. The group-diagnosis results arrange the location of concepts and lacking links at the same location with the goal map's location. An alignment of each excessive link location is central between two connected concepts. The Z-pattern layout is the route of the instructor's eye traveling when they used the location for selecting the proposition in unordering of correctness information. The direction to select the content follows the shape of the letter Z as left to right, top to bottom of visualization screen. It can be used with a hierarchy of concept map that the components are ordered the most important from top to bottom. It can help the instructor to remember the selected- and unselected- excessive links even in the unstructured concept maps. We call this way to provide supplementary instruction as "basic strategy" in this paper.

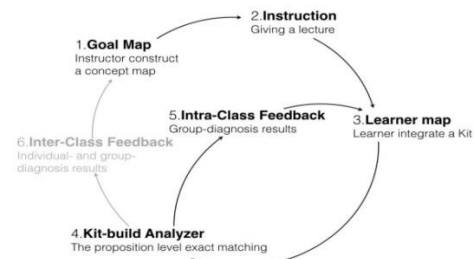


Fig. 10. The arrangement of the KB map on formative assessment.

On the other hand, because the correctness and confidence information are provided in the experiment group, it is assumed that the ordering of supplementary content is different from the ordering of the basic strategy. The difference between the basic strategy and the actual ordering in the practical uses in the experiment group demonstrate the effect of confidence information.

### C. Description of Procedure

The KB map with confidence tagging was utilized in ten science classes. All of the learners were requested to construct the learner map and tagging the confidence two times in each class. The first constructing was requested at the middle of class after the instructor lectured the content, and the second constructing was requested after the instructor gave the supplementary lecture at before the end of class. On the other hand, the different diagnosis results were provided to the instructors for investigating the behavior. A paired class consists of a control class where only the correctness information was visualized and an experiment class where the correctness and confidence information were visualized. Three instructors from three different elementary schools are the participants of the practical uses. An instructor A is the lecturer of fourth-grade who conducted one paired class. An instructor B is the lecturer of fifth-grade that conducted two paired classes, and an instructor C is the sixth-grade lecturer who conducted two paired classes. The instructor lectures the same content in both control- and experiment- classes of each paired class. Fig. 11 displays the practical flow of the paired class to distinguish the different diagnosis results between control- and experiment- group. The correctness information was visualized in both classrooms. The confidence information was blinded as the diagnosis results without confidence in the control classes, while the confidence information was visualized as the diagnosis results with confidence in the experiment classes.

Accordingly, there are no different activities in the learner role, while different information visualizing is the different factor of the instructor role. The different behavior of the same instructor should be observed in each paired class, which is the basic assumption to indicate the relation between the instructor's behavior and the confidence information. The same content of lecturing was conducted with the same instructor, but the supplementary lecturing may be different based on the provided information. The instructor will use the confidence information of learners when s/he accepted the information as the valuable information. In contrast, the behavior of the instructor in the experiment class has a possibility to behave as same as in the control class, even the confidence information was visualized.

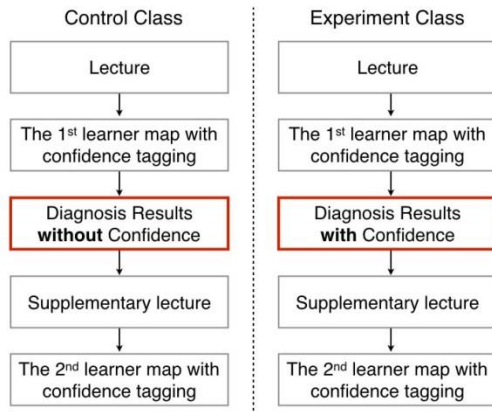


Fig. 11. The practical flows of each paired class.

The primary investigation is about how is the different behavior of the instructors when the system provided the confidence information of their learners. From the assumption, the instructor will use the confidence information for selecting and ordering supplementary content. The gathered evidence of the instructor’s behavior consisted of the order of supplementary content in each class, the discussion session at the end of class, and an information evaluation session of the instructor’s questionnaire. “What is an effect of the different behavior of the instructor?” is analyzed to be three values which contain a normalized learning gain, a discrimination value, and a hit rate. The normalized learning gain of each group was referred to describe the effectiveness of the different behavior of the instructor. The discrimination value illustrates the recognition of the different understanding based on correctness and confidence information. The discrimination value presents how learners have the confidence on the correct proposition and have no confidence on the incorrect proposition. The hit rate focuses only on the correct proposition that learners have confidence. Lastly, the questionnaire was conducted to assess the satisfaction of the KB map with confidence tagging in the aspect of both the learners and the instructors when it was utilized in the classroom situation.

#### IV. RESULTS

##### A. Different Behavior of the Same Instructor

The investigation of the control group is a comparison of excessive links ordering between basic strategy and the actual ordering of each control class, which the assumption is a perfect similarity between the basic strategy and the actual ordering of the class. Fig. 12 shows the goal map of the first paired class. Fig. 13 shows a part of diagnosis results of the control class in the first paired class where the instructor used the filtering function to screen out some excessive links that have the number of the constructor less than three. An observed evidence is the ordering of supplementary content based on the diagnosis results of the class. The first selected excessive link was “composed of 25%,” and supplementary content mentioned to “Water” and “Air” which the action indicates that the most number of excessive links was selected for providing the feedback. The second selected excessive link was “composed of 45%”. These selected excessive links can be ordered by using the correctness information, while the

remaining excessive links have the same tagged number as in ordering of correctness information. The supplementary lecture mentioned to “Water” again with the explanation of “composed of 5%” and the content of “Organic.” Thus, the third selected excessive link was “composed of 5%” on the left-hand side. Then the “composed of 5%” was mentioned with the content of “Organic” again with “Inorganic” content. Hence, the fourth selected excessive link was “composed of 5%” on the right-hand side. The third- and fourth- selected excessive links demonstrate that the location visualization can help the instructor to select the excessive links in unordering of correctness information. Accordingly, the actual ordering of the instructor is the same ordering of basic strategy. The similarity value between basic strategy and actual ordering of the class is 100%. The perfect similarity value illustrates that the instructor used the correctness information and location visualization for ordering feedback, and there are no other factors in this ordering process.

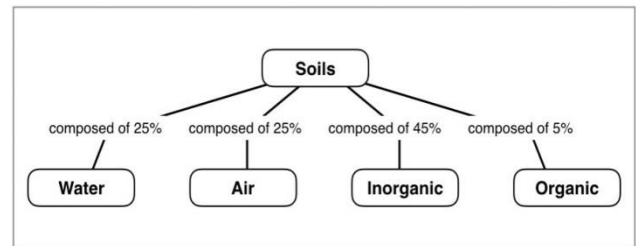


Fig. 12. The goal map of the first paired class.

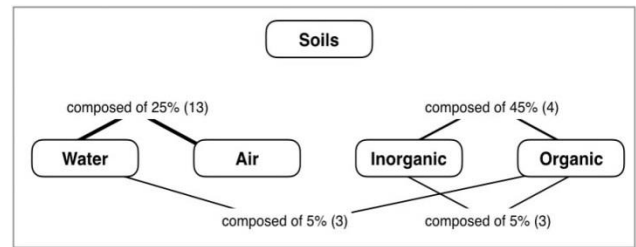


Fig. 13. The group-goal difference map of the control class in the first paired class.

Table I displays the similarity values between the basic strategy and actual ordering of five paired classes. In the control group, all of five control classes can get the perfect similarity value that represents that the instructors used the basic strategy for ordering the supplementary content where the correctness information was provided.

TABLE I. THE PERCENTAGE OF SIMILARITY BETWEEN BASIC STRATEGY AND ACTUAL ORDERING OF FIVE PAIRED CLASSES

Lecturer	Grade of learners	Paired class	Percentage of similarity	
			Control class	Experiment class
Instructor A	4	1 <sup>st</sup> paired	100.00	100.00
Instructor B	5	2 <sup>nd</sup> paired	100.00	60.00
	5	3 <sup>rd</sup> paired	100.00	14.29
Instructor C	6	4 <sup>th</sup> paired	100.00	100.00
	6	5 <sup>th</sup> paired	100.00	16.67

On the other hand, the different order of supplementary content was found in the experiment group where the system provided the correctness and confidence information to the instructor. Imperfect similarity values were found in three of five experiment classes, which indicate the different behavior of the instructors in selecting and ordering the supplementary content.

### B. How the Instructors used the Information of the Diagnosis Results

The different behavior of the same instructor was found when the system provided the different information, and the confidence information has the possibility to encourage the different behavior of the instructor. This section summarizes how the instructors used the diagnosis results from the short discussion sessions with the instructors after finished classes and the evaluation session from the questionnaire of the instructors. The summary mentions to the importance of each information in the diagnosis results, which consist of correctness, confidence information, and location visualization. The instructors commented that the correctness is only one learning evidence in the control group and they focused on the correctness information from the diagnosis results firstly, while the location visualization can help them to point out selected- and remain- excessive links. On the other hand, two learning evidences are provided in the experiment group. The correctness information is still the most important information, and confidence information becomes valuable information as the second priority, then the last priority is visualization location. The result of questionnaire also presents the order of information, which the instructors tried to pay attention to the incorrect proposition first and then looked for its tagged number of confidence information. The incorrect with confidence is the most crucial type of proposition that the all of the instructors want to provide the feedback for this proposition type before the others. Besides, even the strategy of ordering between the control- and experiment- group is different because the different behavior of the instructor on the different diagnosis results, the ordering of the first- and fourth- classes of both groups are the same order with basic strategy as shown in Table I.

Fig. 14 shows an example of the group-goal difference map layout that visualize the group-goal difference map in blinded concept label and linking words for investigating the ordering of the experiment group where the system provides both correctness and confidence information to the instructor. The correctness information is visualized in the form of the number of excessive links for indicating the misunderstanding of learners. The most number of the excessive link is displayed as “Link O (7)” for informing seven learners who connected “Concept A” and “Concept C” with the “Link O.” Thus, the first selected excessive link was selected by using only the correctness information. However, only the correctness information cannot suggest the next selected excessive link because there are six candidates that are possible to be the second selected excessive link. The confidence information is visualized for informing how many learners have the confidence and unconfidence on each excessive link. The tagged number of confidence information on six candidates

suggests that three of three confidences on two excessive links, and two of three confidences on four remaining excessive links. Subsequently, the supplementary lecture mentions to “Link N” with the error explanation, which is according to the “Concept A” and the “Concept D,” and then still keep an attention on the “Link N” again but the error explanation is according to the “Concept C” and “Concept D”. The order of supplementary content demonstrates that the confidence information was used for selecting these selected excessive links. The second selected excessive link is the upper “Link N (3) 3:0”, and the third selected excessive link is the lower “Link N (3) 3:0”. Hence, the order also demonstrates the location visualization was used for ordering when the correctness and confidence information have an equal amount.

Table II displays the used information of ordering process which can represent the amount of time that the instructor used each information. The instructor tended to incorporate the confidence information with the correctness information and location visualization. Thus, we define “CCL” strategy as the ordering supplementary content based on correctness, confidence information, and location visualization respectively. Moreover, there is the possibility, that the instructor used different strategy but both strategies can produce the same order of supplementary content. For instance, the ordering of selected excessive links in the first experiment class was ordered by using five times of correctness and two times of confidence based on CCL strategy. The same ordering can be produced from the basic strategy.

### C. Normalized Learning Gain and Effect Size

The same instructor and the same lecture content are lecturing in each paired class, while the different feedbacks produced the different intervention between the control- and experiment classes. The investigation of normalized learning gains and effect size are presented in this section, and an assumption is the different behavior based on different used strategy affects learning achievements. That means the confidence information effects to the behavior of the instructor, and then the different feedback also effects to the understanding of learners. The normalized learning gain ( $g$ ) is used to represent the effectiveness of the educational intervention [39]. The first learner map was constructed after the instructor gave the lecture (Formative map) and the second learner map was constructed after the instructor gave the supplementary lecture (Final map), which correspond to the arrangement of the KB map on the formative assessment. The learner map scores and the normalized learning gain of each learner can be calculated following:

$$\text{Map score} = \frac{\text{Correct propositions in learner map}}{\text{The number of propositions in the goal map}} \quad (1)$$

$$g = \frac{\text{Final map score} - \text{Formative map score}}{1 - \text{Formative map score}} \quad (2)$$

Correspondingly, the gain of averages ( $\langle g \rangle$ ) was used to indicate the normalized learning gain of class that can be classified into three regions of  $g$  for substantial using following “Low” when ( $\langle g \rangle$ ) less than 0.3, “Medium” when ( $\langle g \rangle$ ) from 0.3 to 0.7, and “High” when ( $\langle g \rangle$ ) more than 0.7 [39].



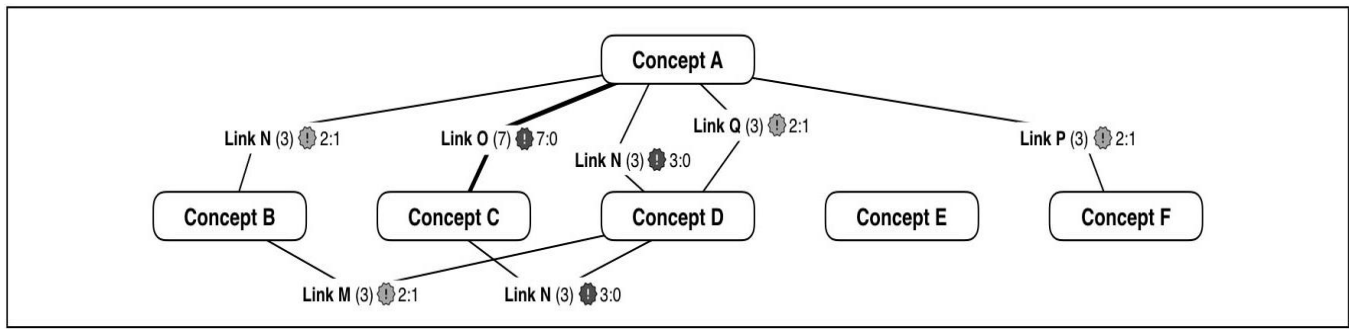


Fig. 14. The layout of the group-goal difference map of the experiment class in the third paired class.

TABLE II. THE INFORMATION USED OF ORDERING IN THE EXPERIMENT GROUP BASED ON CCL STRATEGY

Classroom	Selected excessive links	The number of used time information			Percentage of similarity*
		Correctness	Confidence	Location	
1 <sup>st</sup> experiment class	5	5	2	0	100.00
2 <sup>nd</sup> experiment class	5	5	4	0	60.00
3 <sup>rd</sup> experiment class	7	7	6	6	14.29
4 <sup>th</sup> experiment class	5	5	3	2	100.00
5 <sup>th</sup> experiment class	6	6	6	5	16.67

\*The similarity values of selected excessive links ordering between the basic strategy and CCL strategy

Table III presents the gain of averages and its region of each class. Four experiment classes out of five got better the normalized learning gains than their paired control classes. Especially in the fourth- and fifth- paired classes, there were significant differences in normalized learning gains between experiment class and control class.

Moreover, regarding effect size (Cohen’s *d*) as difference of normalized learning gains between control class and experiment class, they are “large” in the 3<sup>rd</sup> and 5<sup>th</sup> paired classes and they are “medium” in the 4<sup>th</sup> one. There results suggest that the experiment classes were better for learning than control classes.

#### D. Discrimination of the Understanding

The discrimination value ( $d_r$ ) represents the recognition of the difference between what they know and what they do not know [1]. The value is measured based on a proportion of the confident correct proportion and the unconfident incorrect proposition against all of the complete propositions in the learner map. The perfect score indicates the learners are able to discriminate according to an appropriate confidence, which implies the learner has confidence on all of the correct understanding and has no confidence on the misunderstanding.

$$d_r = \frac{\text{Correct with confidence} + \text{Incorrect with unconfidence}}{\text{The number of complete propositions in the learner map}} \quad (3)$$

TABLE III. NORMALIZED LEARNING GAIN OF CLASS AND EFFECT SIZE OF EACH PAIRED CLASS

Paired class	Type of class	Number of learners	$\langle g \rangle$	S.D.	Region of $g$	$d$	$p$ -value <sup>a</sup>
1 <sup>st</sup> paired class	Control class	34	0.57	0.48	Medium	0.23	0.5570
	Experiment class	36	0.67	0.38	Medium		
2 <sup>nd</sup> paired class	Control class	24	0.85	0.46	High	0.13 <sup>b</sup>	0.2660
	Experiment class	26	0.79	0.43	High		
3 <sup>rd</sup> paired class	Control class	25	0.50	0.53	Medium	0.83	0.3019
	Experiment class	25	0.93	0.51	High		
4 <sup>th</sup> paired class	Control class	16	0.29	0.23	Low	0.56	0.0389 <sup>c</sup>
	Experiment class	20	0.47	0.41	Medium		
5 <sup>th</sup> paired class	Control class	17	0.18	0.33	Low	1.49	0.0003 <sup>c</sup>
	Experiment class	20	0.71	0.38	High		

<sup>a</sup>. The  $p$ -value of  $g$  between control- and experiment- class of each paired class.

<sup>b</sup>. The value presents  $|d|$  when the control class has the  $\langle g \rangle$  more than the experiment class, which produces a negative value of  $d$ .

<sup>c</sup>. Statistically significant difference

TABLE IV. AN AVERAGE OF THE DISCRIMINATION VALUE

Group (N=10)	Formative map	Final map	p-value
Control group (5 classes)	0.6007	0.7624	$p < 0.01$
Experiment group (5 classes)	0.6820	0.8842	$p < 0.01$
p-value	0.0794	$p < 0.01$	

Table IV shows the discrimination value of learners and the significant difference between the control group and the experiment group. There was no significant difference between the formative map of the control- and experiment- group ( $p=0.794$ ), which means that the learners have an ability to discriminate about their knowledge not much different after lecturing. The feedback of instructors improved discrimination of learners in both groups significantly ( $p<0.01$ ). Then, there was a significant difference of final map between the control- and experiment- group ( $p<0.01$ ). These results suggest that the correctness and confidence based feedback can improve the discrimination of their confidence on their understanding better than the correctness based feedback.

### E. Certainty of the Understanding

The confidence on the incorrect proposition is the worst situation that the instructors attempt to correct those misunderstanding by providing the supplementary lecture based the diagnosis results. On the other hand, the confidence on the correct proposition is the best situation for representing the certainty of the understanding. The hit rate (HR) represents consistency with the interpretation that if a correct response is covertly selected, then its execution helps the learner to confirm its correctness [1]. The value is measured based on a proportion of the number of confident correct propositions against the number of correct propositions in the learner map.

$$HR = \frac{\text{Correct proposition with confidence}}{\text{The number of correct propositions in the learner map}} \quad (4)$$

Table V shows the hit rate and the significant difference between two learner maps of two groups. There was no significant different between control- and experiment- group ( $p=0.1976$ ) that means learners have not much different confidence on the correct answers after lecturing. Then the feedback of instructors can improve confidence on the correct answers in both groups significantly ( $p<0.01$ ). There was also a significant difference of final map between the control- and experiment- group ( $p<0.05$ ), which suggests that the correctness and confidence based feedback can improve the certainty of the understanding better than the correctness based feedback.

TABLE V. AN AVERAGE OF THE HIT RATE

Group (N=10)	Formative map	Final map	p-value
Control group (5 classes)	0.7430	0.8888	$p < 0.01$
Experiment group (5 classes)	0.6714	0.9587	$p < 0.01$
p-value	0.1976	$p < 0.05$	

### F. Changing of Proposition based on the Confidence

For more emphasis on the confidence of learners, Table VI shows a possibility of proposition changing based on the confidence information from the formative map to the final map. The analysis of change of proposition type presents that the propositions with unconfidence are easier to change than the propositions with confidence. Particularly, the changing of unconfidence propositions to confident correct propositions of experiment group is 80.30%, while 69.60% unconfidence propositions of the control group are changed to confident correct propositions. The proposition changing suggests that the correctness and confidence based feedback can help the learners to improve their understanding and get more confidence better than the correctness based feedback.

TABLE VI. A PROPOSITION CHANGING BASED ON THE CONFIDENCE OF LEARNERS FROM THE FORMATIVE MAP TO THE FINAL MAP

Group (N=10)	Percentage of proposition changing	
	Confidence	Unconfidence
Control group (5 classes)	33.07%	66.97%
Experiment group (5 classes)	33.08%	85.40%

### G. Results of the Questionnaire

The questionnaire was conducted to the learners who participated in the practical uses, which content of the questionnaire contains three sessions following the overview of the KB map with confidence tagging, emphasizing on the effect of confidence tagging, and the effect of instructor’s feedback. Fig. 15 displays a part of the questionnaire of learners. The positive evaluations received from the learners by the questionnaire. Such as the first questions, 60.70% of learners “strongly agree” enjoy constructing the learner map and tagging of the confidence. 51.26% “strongly agree” and 29.14% “agree” are the results of the second question about constructing the map and tagging confidence are useful for expressing the understanding of lecture content. The confidence tagging as an additional task did not disturb the learners in the structuring task, which 34.67% and 31.66% “strongly agree” and “agree” on they feel free to tagging their confidence respectively as the results of the fourth question. Finally, the results of seventh- and eighth- questions have more than fifty percent on “strongly agree” that the instructor’s feedback in the form of the supplementary lecture can help learners to get more understanding and get more confidence. The results of learner’s questionnaire illustrate the satisfaction of learners that suggests that the learners accepted the mechanism of the KB map with confidence tagging.

The questionnaire of the instructor was also conducted for investigating the aspect of the instructors when the KB map was utilized in their lecture classes. Fig. 16 displays a part of the instructor’s questionnaire. The results of the questionnaire demonstrate the positive satisfaction of the instructors.

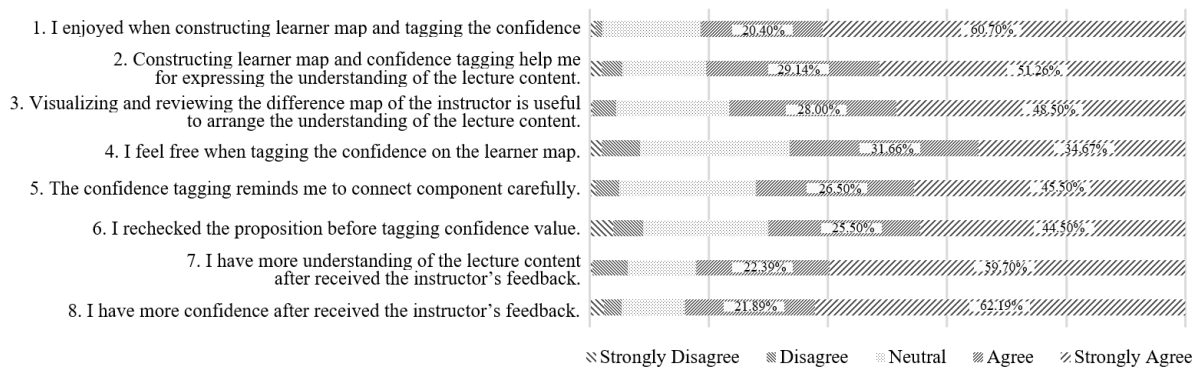


Fig. 15. A part of learner's questionnaire and its results.

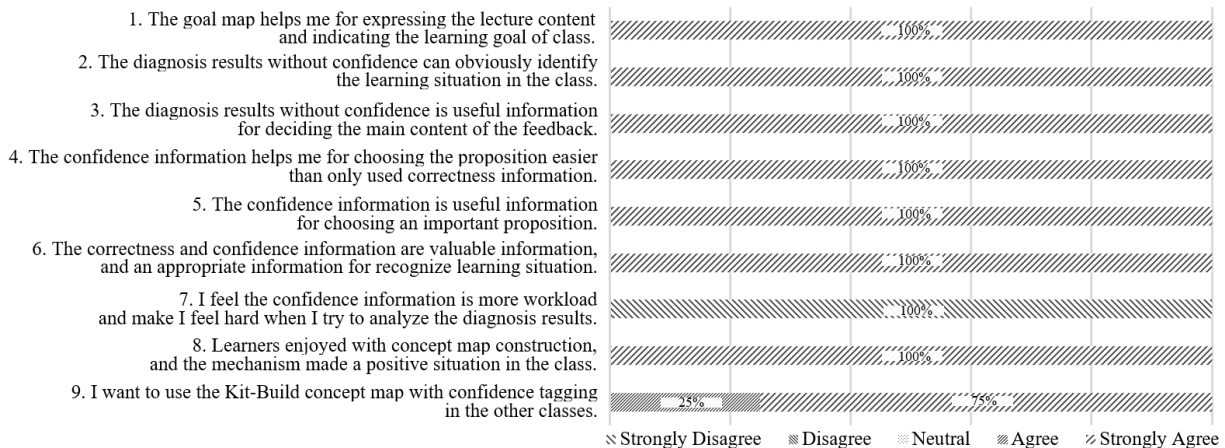


Fig. 16. A part of instructor's questionnaire and its results.

The goal map creating can help the instructors to express the lecture content, and indicate the learning goal as the result of the first question. The results from the second- to sixth-questions present that all instructors gave “strongly agree” to the diagnosis results, which are useful information for visualizing the current learning situation, identifying the critical misunderstanding of learners, until selecting and ordering the supplementary content. Moreover, the instructors also strongly agreed on the eighth question that their learners enjoyed with the mechanism which formed the positive environment for the learning situation. On the other hand, the instructor gave “strongly disagree” on the seventh question that the confidence information was more workload when analyzing the diagnosis results. Thus, the instructor accepted the diagnosis results that include the correctness and confidence information. Notably, the diagnosis results with the confidence information are useful information for selecting and ordering the supplementary feedback, which is more satisfactory than no confidence information.

## V. DISCUSSION

In this study, we present the encouragement of correctness and confidence information with the KB map with confidence tagging for selecting and ordering the supplementary content as the feedback of the instructors in the lecture classes. The KB map creates an opportunity for an instructor to assess a current learning situation, which the instructor can give the feedback to learners for improving learning achievements in the class

period. The different behavior of the instructors was observed when the system provided only the correctness information in the control group, while the correctness and confidence information were provided in the experiment group. The ordering of the supplementary content demonstrates how the instructor used the correctness, confidence information, and location visualization.

The observed evidence of the practical uses can represent the relation between the instructor's behavior and the confidence information of learners. The instructors did not only use the confidence information in selecting and ordering the supplementary content, but we also found the mentioning to the confidence of learners on some selected excessive links in the supplementary lecture of the experiment group when the instructor received the confidence information. Correspondingly, the relation of instructor's behavior and learning evidence suggests that the different behavior of the instructors is positive changing to improve the learning achievements and also improve the confidence of learners. The normalized learning gain of class ( $\langle g \rangle$ ) and effect size (Cohen's  $d$ ) illustrate that the correctness and confidence based feedback of the experiment group is more effective than the only correctness based feedback of the control group. The discrimination value ( $d_r$ ) demonstrates that the learners of experiment group can discriminate the different understanding based on correctness and confidence better than the learners of control group significantly. Similarly, the hit rate ( $HR$ ) shows that the learners of experiment group have an ability to

REFERENCES

represents consistency with the interpretation better than the learners of control group significantly. These results of the practical uses suggest that the confidence information of learners affects the instructor's behavior and then the different behavior of the instructor effects to the learning achievements continuously. In addition, the results of questionnaire present the positive satisfaction of both instructors and learners when the KB map with confidence tagging was utilized in the lecture classes. The learners accepted the mechanism for representing their understanding and their confidence. The instructors accepted that the confidence information of learners is valuable information for recognizing the learning situation. Nevertheless, the content details of the supplementary lecture were not investigated in this experiment such as what kind of feedback was designed from only correctness, or correctness and confidence information.

VI. CONCLUSION

Even the correctness assessment can determine the knowledge of learners, the quality of that knowledge cannot be identified by using only the correctness information. We propose the KB map with confidence tagging that can provide the mechanism to learners for representing their understanding and identifying their confidence on their understanding. The learner map and confidence of each proposition are the learning evidence, which the learner map can represent the understanding of learners in the lecture content and the confidence tagging promotes them to reconsider their propositions again. The system facilitates learners to create learning evidence in a class period and identify the current learning situation through diagnosis results immediately. Subsequently, the learning evidence of learners affects the instructor behavior directly when they accepted the information as a valuable information. The supplementary lecture based on the correctness and confidence information is utilized as evidence-based feedback of the instructor, which is a key of formative assessment to improve learning achievements in the classroom situations.

Moreover, the different behavior of the same instructor illustrates the utilizing of the confidence information on the supplementary lecture that can demonstrate that the instructor accepted the confidence information as the valuable information. The confidence information can encourage the strategy for selecting and ordering the supplementary content. The results of the practical uses suggest that the different feedback of the instructor is important through normalized learning gains and effect size, which the correctness and confidence based feedback can improve the learning achievements and confidence of learners concurrently.

For the future work, the individual feedback will be focused based on the current ability of the KB map with confidence tagging. Even the instructor can improve the learners understanding, some propositions are disregarded such as the correct proposition with unconfidence. Consequently, we aim to direct to all learners and support all their propositions via the KB map with confidence tagging for improving the learning achievements in the form of system feedback.

- [1] D. P. Hunt, "The concept of knowledge and how to measure it," *Journal of intellectual capital*, vol. 4, no. 1, pp. 100-113, 2003.
- [2] W. Bruine de Bruin, A. M. Parker, and B. Fischhoff, "Individual differences in adult decision-making competence," *Journal of personality and social psychology*, vol. 92, no. 5, pp. 938-956, 2007.
- [3] A. Efklides, and A. Tsiora, "Metacognitive experiences, self-concept, and self-regulation," *Psychologia*, vol 45, no. 5, pp. 222-236, 2002.
- [4] A. Efklides, "Metacognition and affect: What can metacognitive experiences tell us about the learning process?," *Educational research review*, vol. 1, no. 1, pp. 3-14, 2006.
- [5] S. Kleitman, and T. Moscrop, "Self-confidence and academic achievements in primary-school children: Their relationships and links to parental bonds, intelligence, age, and gender," In *Trends and prospects in metacognition research*, Springer US, pp. 293-326, 2010.
- [6] S. Kleitman, L. Stankov, C. M. Allwood, S. Young, and K. K. L. Mak, "Metacognitive self-confidence in schoolaged children," In *Self-directed learning oriented assessments in the Asia-Pacific*, Springer Netherlands, pp. 139-153, 2012.
- [7] L. Stankov, and J. Lee, "Confidence and cognitive test performance," *Journal of Educational Psychology*, vol. 100, pp. 961-976, 2008.
- [8] L. Stankov, J. Lee, and I. Paek, "Realism of confidence judgments," *European Journal of Psychological Assessment*, vol. 25, no. 2, pp. 123-130, 2009.
- [9] G. Heron, and J. Lerpiniere, "Re-engineering the multiple choice question exam for social work," *European Journal of Social Work*, vol. 16, no. 4, pp. 521-535, 2013.
- [10] S. M. Cisar, P. Cisar, and R. Pinter, "True/false questions analysis using computerized certainty based marking tests," In *Intelligent Systems and Informatics, 7th International Symposium on*, pp. 171-174, 2009.
- [11] T. Hirashima, K. Yamasaki, H. Fukuda, and H. Funaoi, "Framework of kit-build concept map for automatic diagnosis and its preliminary use," *Research and Practice in Technology Enhanced Learning*, vol. 10, no. 17, pp. 1-21, 2015.
- [12] K. Yoshida, K. Sugihara, Y. Nino, M. Shida, and T. Hirashima, "Practical Use of Kit-Build Concept Map System for Formative Assessment of Learners' Comprehension in a Lecture," *Proc. of ICCE2013*, pp. 892-901, 2013.
- [13] K. Yoshida, T. Osada, K. Sugihara, Y. Nino, M. Shida, and T. Hirashima, "Instantaneous Assessment of Learners' Comprehension for Lecture by Using Kit-Build Concept Map System," In *International Conference on Human Interface and the Management of Information*, Springer, Berlin, Heidelberg, pp. 175-181, 2013.
- [14] J. Pailai, W. Wunnasri, K. Yoshida, Y. Hayashi, and T. Hirashima, "The practical use of Kit-Build concept map on formative assessment," *Research and Practice in Technology Enhanced Learning*, vol. 12, no. 20, pp. 1-23, 2017.
- [15] J. E. Bruno, "Using testing to provide feedback to support instruction: A reexamination of the role of assessment in educational organizations," In *Item banking: Interactive testing and self-assessment*, Springer, Berlin, Heidelberg, pp. 190-209, 1993.
- [16] A. R. Gardner-Medwin, and M. Gahan, "Formative and summative confidence-based assessment," *Proc. of the 7th CAA Conference*, Loughborough: Loughborough University, pp. 147-155, 2003.
- [17] T. Gardner-Medwin, and N. Curtin, "Certainty-based marking (CBM) for reflective learning and proper knowledge assessment," In *REAP Int. Online Conf. on Assessment Design for Learner Responsibility*, 2007.
- [18] G. Yuen-Reed, and K. B. Reed, "Engineering Student Self-Assessment through Confidence-Based Scoring," *Advances in Engineering Education*, vol. 4, no. 4, pp. 1-23, 2015.
- [19] A. R. Gardner-Medwin, "Optimisation of certainty-based assessment scores," 2013, *Proc. of The Physiological Society, The Physiological Society*.
- [20] C. M. Moss, and S. M. Brookhart, *Advancing formative assessment in every classroom: A guide for instructional leaders*, ASCD, 2010.
- [21] D. Wiliam, C. Lee, C. Harrison, and P. Black, "Teachers developing assessment for learning: Impact on student achievement," *Assessment in Education: Principles, Policy & Practice*, vol. 11, no. 1, pp. 49-65, 2004.

- [22] R. Ballantyne, K. Hughes, and A. Mylonas, "Developing procedures for implementing peer assessment in large classes using an action research process," *Assessment & Evaluation in Higher Education*, vol. 27, no. 5, pp. 427-441, 2002.
- [23] J. D. Novak, and A. J. Cañas, "The theory underlying concept maps and how to construct and use them", Technical Report IHMC CmapTools, 2008.
- [24] M. Buldu, and N. Buldu, "Concept mapping as a formative assessment in college classrooms: Measuring usefulness and student satisfaction," *Procedia-Social and Behavioral Sciences*, vol. 2, no. 2, pp. 2099-2104, 2010.
- [25] D. L. Trumpower, and G. S. Sarwar, "Formative structural assessment: Using concept maps as assessment for learning," *Proc. of the Fourth International Conference on Concept Mapping*, vol. 22, pp 132-136, 2010.
- [26] J. Schacter, H. E. Herl, G. K. W. K. Chung, R. A. Dennis, and H. F. O'Neil, "Computer-based performance assessments: A solution to the narrow measurement and reporting of problem-solving," *Computers in Human Behavior*, vol. 15, no. 3, pp. 403-418, 1999.
- [27] I. L. G. Hsieh, and H. F. O'Neil, "Types of feedback in a computer-based collaborative problem-solving group task," *Computers in Human Behavior*, vol. 18, no. 6, pp. 699-715, 2002.
- [28] C. C. Chiou, "The effect of concept mapping on students' learning achievements and interests," *Innovations in Education and Teaching International*, vol. 45, no. 4, pp. 375-387, 2008.
- [29] P. Chularut, and T. K. DeBacker, "The influence of concept mapping on achievement, self-regulation, and self-efficacy in students of English as a second language," *Contemporary Educational Psychology*, vol. 29, no. 3, pp. 248-263, 2004.
- [30] N. Aghakhani, H. S. Nia, S. Eghtedar, and C. Torabizadeh, "The Effect of Concept Mapping on the Learning Levels of Students in Taking the Course of," *Jundishapur Journal of Chronic Disease Care*, vol. 4, no. 2, pp. 1-5, 2015.
- [31] J. Pailai, W. Wunnasri, Y. Hayashi, and T. Hirashima, "Ongoing Formative Assessment with Concept Map in Proposition Level Exact Matching," *Proc. of ICCE2016*, pp. 79-81, 2016.
- [32] J. Pailai, W. Wunnasri, Y. Hayashi, and T. Hirashima, "Automatic Concept Map Assessment in Formative Assessment Approach," *Engineering Research Council and the 76th Advanced Learning Science Society for Artificial Intelligence*, vol. 5, no. 3, pp. 48-53, 2016.
- [33] E. M. Taricani, and R. B. Clariana, "A technique for automatically scoring open-ended concept maps," *Educational Technology Research and Development*, vol. 54, no. 1, pp. 65-82, 2006.
- [34] B. S. Bloom, M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl, "Taxonomy of educational objectives, handbook I: The cognitive domain," 1956.
- [35] M. Alkhateeb, Y. Hayashi, T. Rajab, and T. Hirashima, "Comparison between kit-build and scratch-build concept mapping methods in supporting EFL reading comprehension," *The Journal of Information and Systems in Education*, vol. 14, no. 1, pp. 13-27, 2015.
- [36] M. Alkhateeb, Y. Hayashi, T. Rajab, and T. Hirashima, "Experimental Evaluation of the KB-mapping Method to Avoid Sentence-by-Sentence Map-building Style in EFL Reading with Concept Mapping," *The Journal of Information and Systems in Education*, vol. 15, no. 1, pp. 1-14, 2016.
- [37] M. Alkhateeb, Y. Hayashi, T. Rajab, and T. Hirashima, "Experimental Use of Kit-Build Concept Map System to Support Reading Comprehension of EFL in Comparing with Selective Underlining Strategy," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 4, pp. 80-87, 2016.
- [38] H. Funaoi, K. Ishida, and T. Hirashima, "Comparison of kit-build and scratch-build concept mapping methods on memory retention," *Proc. of ICCE2011*, pp. 539-546, 2011.
- [39] R. R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *American journal of Physics*, vol. 66, no. 1, pp. 64-74, 1998.