広島大学学術情報リポジトリ
Hiroshima University Institutional Repository

Title	Reliability Investigation of Automatic Assessment of Learner-Build Concept Map with Kit-Build Method by Comparing with Manual Methods
Auther(s)	Wunnasri, Warunya; Pailai, Jaruwat; Hayashi, Yusuke; Hirashima, Tsukasa
Citation	Lecture Notes in Computer Science , 10331 : 418 - 429
Issue Date	2017
DOI	10.1007/978-3-319-61425-0_35
Self DOI	
URL	http://ir.lib.hiroshima-u.ac.jp/00046326
Right	The final authenticated version is available online at https://doi.org/10.1007/978-3-319-61425-0_35. This is not the published version. Please cite only the published version. この論文は出版社版でありません。引用の際には出版社版をご確認ご利用ください。
Relation	

Hiroshima University Institutional Repository

Reliability Investigation of Automatic Assessment of Learner-Build Concept Map with Kit-Build Method by Comparing with Manual Methods

Warunya Wunnasri, Jaruwat Pailai, Yusuke Hayashi & Tsukasa Hirashima

Graduate school of Engineering, Hiroshima University, Japan

warunya,jaruwat,hayashi,tsukasa@lel.hiroshima-u.ac.jp

Abstract. This paper describes an investigation into the reliability of an automatic assessment method of the learner-build concept map by comparing it with two well-known manual methods. We have previously proposed the Kit-Build (KB) concept map framework where a learner builds a concept map by using only a provided set of components, known as the set "kit". In this framework, instant and automatic assessment of a learner-build concept map has been realized. We call this assessment method the "Kit-Build method" (KB method). The framework and assessment method have already been practically used in classrooms in various schools. As an investigation of the reliability of this method, we have conducted an experiment to compare the assessment results of the method with the assessment results of two other manual assessment methods. In this experiment, 22 university students attended as subjects and four as raters. It was found that the scores of the KB method had a very strong correlation with the scores of the other manual methods. The results suggest that automatic assessment of the Kit-Build concept map can attain almost the same level of reliability as well-known manual assessment methods.

Keywords: Concept Map Assessment Method, Kit-Build Concept Map, Reliability

1 Introduction

Concept maps were developed in 1972 in Novak and Musonda's research program [1] which investigated changes in children's knowledge of science. Novak and Musonda's research was based on the learning psychology of Ausubel et al. [2] which discussed the assimilation of new knowledge into existing knowledge by learners. A concept map represents conceptual understanding via connections and links between concepts. A concept in a concept map can be a term or symbol that is enclosed in a box, and a link is a line that is connected to two concepts. A linking word is a word on the link that represents the relationship between concepts. To build the concept map, creators have to organize their knowledge following their target. They can limit the scope of their concept map by constructing a concept map for answering the focus

question. Then the creators build a concept list from the main idea of the content and they order these concepts from general to more specific aiding in hierarchical construction. Proposition of the concept map, or unit of meaning, can be constructed from linking two or more concepts via a proper relationship. The concepts should be ordered by placing the general concept in the top hierarchy and specific concepts at the bottom [3]. Moreover, concept maps can help learners to significantly reduce their learning cognitive load, because concept maps assist in the integration of knowledge and facilitate learners in their independent learning and thinking [4]. Due to these characteristics, concept map is used to organize and represent knowledge extensively.

Afterward, the concept maps are used in a classroom situation for checking learners' understanding. Several educational researchers proposed the concept map assessment method for checking learners' understanding. These assessment methods, which are processed manually, are reasonable for evaluating concept maps, but they entail high costs, such as time and human workload, for scoring each concept map. Hence, an automatic concept map assessment is proposed for decreasing time cost and human workload.

The Kit-Build concept map (KB map) is a framework to realize automatic concept map assessment [5, 6]. In the KB map framework, a learner builds a concept map by using only a provided set of components, referred to as the set "kit". Instant and automatic assessment of a learner-build concept map, realized in this framework, is referred to as the "Kit-Build method" (KB method). In this framework, the set of components are made by decomposing a concept map that is built by a responsible teacher. This map is called the "teacher-build map". The responsible teacher is requested to build the teacher-build map as a criterion to assess a learner's comprehension for a specific topic or teaching. Then, a learner is requested to build a concept map to express his/her comprehension for the same topic or teaching. Because all components of the learner-build map are the same as the teacher-build map, automatic assessment of a learner-build map is realized by comparing the learner-build map with the teacher-build map. KB map and assessment methods have already been practically used in classrooms in various schools, for example, in science learning in elementary schools [7, 8], geography in junior high schools [9], and the learning of English as a second language [10].

These practices have shown that the KB map is suitable for use in teaching situation where the instructor gives directions followed by instructor's interpretation. However, we have not previously compared the KB method with other well-known manual methods that are accepted as reliable. Although the automatic assessment method has advantages over manual assessment, for example, real time assessment/feedback, load reduction of the rater/teacher, etc., the reliability of automatic assessment requires investigation. In this study, the results of manual methods were assumed to be reliable and we compare the assessment results of the KB method with the assessment results of two other manual assessment methods. As the two manual methods, (1) structural scoring proposed by Novak [11] and (2) relational scoring proposed by McClure [12] were adopted. We conducted an experiment where 22 university students were designated as subjects and four were designated as raters. The results of the experiment showed that the scores of the KB method had a statisti-

cally significant correlation with the scores of the other manual methods. The results suggest that automatic assessment using the KB method can attain almost the same level of reliability as well-known manual methods.

2 Overview of the Concept Map Assessment Method

2.1 Manual Concept Map Assessment Methods

A manual concept map assessment method is used by a human who can understand the meaning of words in the concept map well. The human is often called a "rater". In this study, we focus on the methods that pay attention to the structure of a concept map and the meaning of the proposition of a concept map.

Several concept map assessment methods evaluate the concept map by investigating the structure of the map, such as, the levels of the hierarchy, the characteristics of the branch, etc. In this study, we focus on the structural scoring of Novak and Gowin [11] as a typical structural method. This method gives high scores for each correct level of the hierarchy and each valid crosslink because ordering the concepts into the hierarchy, and connecting the crosslinks, can facilitate the constructor's creative thinking. However, structural scoring, which tends to score the structure more than the meaning, may be the cause of substantial meaning-leakage in a concept map.

Many manual assessment methods which pay more attention to the meaning of a proposition for scoring the concept map, rather than the structure, have been proposed. They focus on language and understanding of the representation. These meaningful methods always have a printed set of criteria as the rubric for assessing knowledge and giving feedback. From investigating various meaning methods, we focused on the relational scoring from McClure and Bell [12], which is referred to as relational scoring in this paper, and is a common concept map assessment method. In the study by McClure et al. [13], they requested 63 students to construct concept maps by using 20 provided concepts, creating their own linking words. Then, 12 raters scored individual maps by assessing each proposition on the concept map separately. The rater awarded scores of zero to three points for each proposition based on the suitability of the meaning of the proposition. The authors claimed that this method has the highest reliability when using the criteria map, (teacher-build map), using the holistic method and the structural method as comparisons (Novak and Gowin structural scoring). The authors confirmed this result by using the g-coefficient value. Based on these considerations, we have designed an experiment for testing the reliability of a manual method, similar to the experiment of McClure et al. We have selected the structural scoring proposed by Novak and Gowin, and the relational scoring proposed by McClure and Bell, to compare with the KB map proposed in the current study.

2.2   Kit-Build Concept Map and Automatic Assessment

The Kit-Build concept map framework is one of the automatic concept map assessment methods that use a teacher-build map to compare with the learner-build map by using exact matching at the propositional level. It is utilized in the form of a learning task or exercise for checking learners' comprehension of a topic that they have already learned. The task of the KB map is separated into two subtasks. The first is the segmentation task where a teacher is requested to prepare the teacher-build map, which is an expression of an eligible comprehension of the topic for the teacher. An example of the teacher-build map is illustrated in Figure 1. After submitting the teacher-build map to the server, the teacher-build map is extracted to be the kit that contains a list of concepts and relationships from the teacher-build map. The kit from the teacher-build map in Figure 1 is shown in Figure 2. Moreover, this kit is provided to help learners to reduce their cognitive load more than the traditional concept map, where they must create all components themselves. Using the kit, the learners are required only to recognize the components.
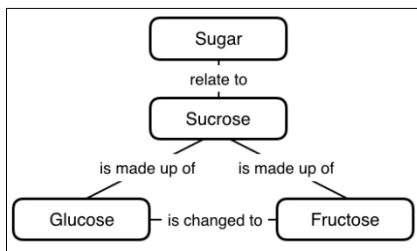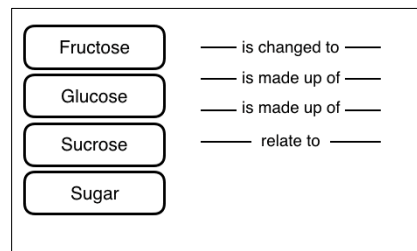

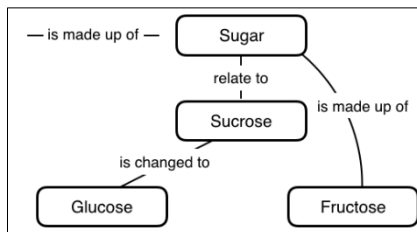**Fig. 1.** Teacher-build map

Fig. 2. Kit
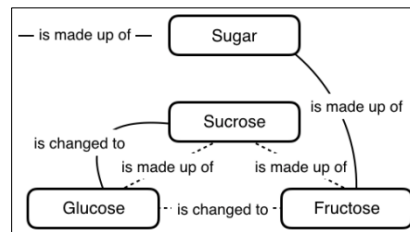

Fig. 3. Learner-build map

Fig. 4. Difference map

The second task is called the structuring task. Learners are given the learning task of reconstructing a concept map by using the kit, creating a map which is referred to as the learner-build map (Figure 3). After the learner-build maps are uploaded to the server, the KB map will evaluate learner-build maps by exactly matching each learner's proposition with the teacher-build map's proposition. For example, the relationship between the concepts "Sugar" and "Sucrose" is checked. If the relationship is identified as "related to," the score for this learner-build map will increase by one point. In the case of the concepts "Sucrose" and "Glucose," if the learner connected

them by using the relationship "is changed to," this does not exist in the teacher-build map. Following the teacher-build map, the relationship of this proposition should be "is made up of", so this proposition is not awarded any point from the system. This corresponds to the scoring by propositional level exact matching method. This method makes the KB map different from the manual methods which allow learners to create their own linking words, preventing the learner-build map from being straightforwardly compared with the criteria map. The manual methods require time for considering the meaning of each proposition carefully. After checking the connections of the learner-build maps by the propositional level exact matching, the system will generate a score in a percentage format. The instructor can also investigate learners' misunderstanding individually as a difference map and can find the overview of all learners by overlaying all learner-build map, as the group map, and the group-goal difference map on the analysis screen of the KB. In the difference map, three types of error link are represented as shown in Figure 4. The lacking link, which is represented by a dashed line, is a link that exists in the teacher-build map but does not exist in learner-build map. The excessive link, which is shown as a solid line, is a link that occurs in learner-build map but does not occur in the teacher-build map. Lastly, a solid line that is not connected to any concepts in the learner-build map is the leaving link. The instructor can use these links to find the holistic leaking understanding of all learners. Following the KB map framework's ability, the instructor can use the KB map to check understanding of individuals or groups of learners, and can use the diagnosis result to discuss with learners the meaning of each of the error links. After error link analysis, the instructor can adjust the teacher-build map or teach learners about the content that learners have not understand completely.

In the other automatic methods, only a list of concepts is provided to learners. Because these methods also allow learners to create their own linking words, they cannot compare a learner-build map with the criteria map straightforwardly. Hence, they require synonym word matching, which is very flexible for evaluation using the meaning of words, but which has not yet reached a sufficient level of accuracy. In contrast, the KB map provides the kit which can be assessed by using the propositional level exact matching and can create informative diagnosis results. Moreover, the KB map can provide the group map and group-goal difference map, which can support the instructor in analyzing comprehension in both an individual learner and as an overview of the whole class. These are the prominent advantages of the KB map when it is utilized in a classroom situation.

3   Research Methodology

To confirm the reliability of the KB map, we designed an experimental procedure to compare the KB map and the manual methods in terms of their ability to assess the comprehension of learner on a topic. Usually, the KB map is used in teaching situations, however, it is desirable to ensure that the KB map as can be used in a reading situation also. Hence, the experiment was designed to operate in two learning situations. Moreover, to compare the difference between the KB map and the manual

method, the important attributes of the concept map assessment method are shown in Table. 1.

Two typical scoring methods, which are widely used for assessing concept maps, namely the structural scoring as structural level analysis, and the relational scoring as propositional level analysis, were chosen for comparison. The manual method is inferred from the research of McClure et al. [13], who provided a list of concepts to learners and requested that they construct concept maps by creating linking words themselves. The synonym matching method was used for evaluating the meaning of each proposition. However, the KB map provides both the concepts and the linking words, which are decomposed from the teacher-build map, to learners. Thus, the automatic exact matching method can be used for checking the correctness of each proposition.

Table 1. Comparison between the attributes of each concept map assessment method

Assessment Method	Assessment			Provided Items	
	Raters	Level of Analysis	Matching Method	Concepts Words	Linking Words
Structural Scoring	Manual	Structural	Synonym	Provided	Not Provided
Relational Scoring	Manual	Propositional	Synonym	Provided	Not Provided
Kit-Build Concept Map	Automate	Propositional	Exact	Provided	Provided

3.1 Subjects

Subjects for this study were recruited from university students who possessed a good level of English. The 22 students, who were volunteers from various education fields, were given the role of learners. They were given introductory training in concept maps before participating in the experiment. Four students, who were familiar with the use of the concept map and understood the content of the experiment material well, were assigned as raters. These raters were given an explanation of the procedure of each assessment method, and they were required to study the procedures carefully before scoring the learner-build map. In addition, one graduate student was assigned the role of instructor. The instructor was required to prepare the article and teaching material for the experiment and the instructor was also required to construct the teacher-build map following specific instructions. In this study, the article "Sugar", which uses common explanatory words, was chosen for the learning process. This article contained three sections, each covering one third of a page, defined as the introduction to sugar, types of sugar and how sugar is produced [14].

3.2 Map Production

Initially, the instructor chose a 1,594 word article, prepared the teaching materials and built the teacher-build map. The teacher-build map contained 15 concepts and 16 relationships. In the study, learners were requested to read the article in ten minutes, and they were then provided with the list of concepts. Next, they were required to

create linking words by themselves for the construction of a concept map in 15 minutes using the CMapCloud application [15], as illustrated in Figure 5. These learner-build maps were scored by the two manual methods. The learners were then asked to construct a concept map again in 15 minutes by integrating the kit of the KB map, which provided both a list of concepts and a list of linking words. The initial representation of the KB map in this experiment is shown in Figure 6. After the learners had completely connected the propositions and uploaded their map to the server, these learner-build maps were evaluated using the KB map assessment method based on exact matching at the propositional level.

After the reading session concluded, the instructor taught learners based on the same reading article but following the instructor's interpretation using 16 slides delivered over ten minutes. Afterward, learners were required to construct the learner-build maps following the same procedure as in the reading situation, namely, constructing learner-build maps by creating linking words by themselves and integrating the kit to create a learner-build map using the KB map. When learners completed all learner-build map construction, they were asked to answer a questionnaire. The procedure of this experiment is displayed in Figure 7.
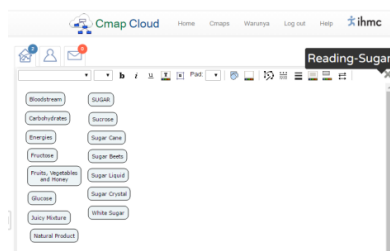

Fig. 5. CMapCloud screen
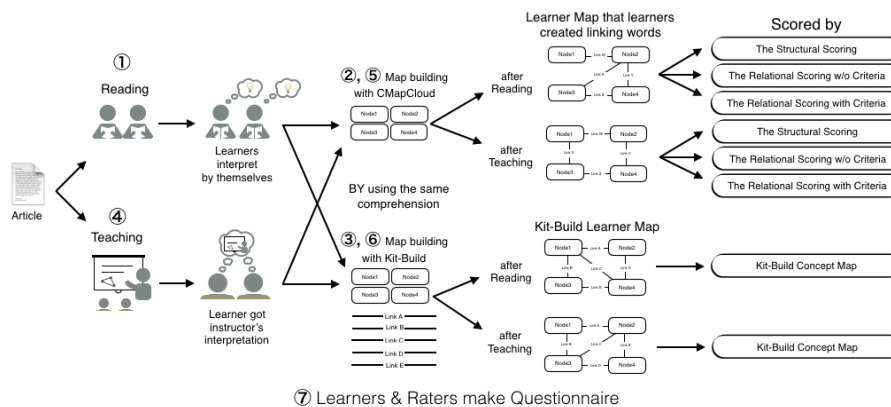

Fig. 6. Kit-Build concept map screen

Fig. 7. Experiment procedure

3.3 Concept Map Scoring by Manual Methods

The concept maps, which were constructed using CMapCloud, were scored by three manual methods that contained, (a) the Novak and Gowin structural scoring (the structural scoring), (b) the McClure and Bell relational scoring without the criteria map (the relational scoring without criteria map) and (c) the McClure and Bell relational scoring with criteria map (the relational scoring with criteria map). The raters were required to read the instructions of each assessment method carefully without time restrictions. The score of the manual methods was normalized to a percentage score by using the perfect score for each method. After the scoring was completed, the raters were requested to complete the questionnaire. Procedures for each method were prepared based on the description in [13]. The reliability of the results of the manual methods is discussed in Section 4.

3.4 Questionnaires

The questionnaires were assigned to both raters and learners to assess their familiarity with concept mapping and their opinion of the experiment. For the raters, the questionnaire contained two parts. The first part of the questionnaire assessed their familiarity with the concept map and with the content of the article. The questionnaire also asked about their disposition when they were scoring the concept maps. The second part of the questionnaire requested raters to rank each scoring method in four aspects covering (i) hardness of decision, (ii) use of memory, (iii) time taken and (iv) reasonableness of the score.

For the learners' questionnaires, the aim was to assess their background in concept mapping and in the content of the article. A further aim was to understand how their experiences differed when constructing the concept maps by creating their own linking words and when using the KB map.

4 Experimental Results and Discussion

4.1 Correspondence of the KB Map and the Manual Method

The reliability of the manual method.
To confirm the KB map's reliability as a framework for assessing learners' comprehension of a topic by comparing with reliable manual methods, we aim to first investigate the reliability of the manual methods. The scores from three manual methods: (a) the structural scoring, (b) the relational scoring without criteria map and (c) the relational scoring with criteria map, were used to perform generalizability analysis through the GNOVA software [16] which returns the g-coefficient, as used in the reliability investigation by McClure et al. [13]. The g-coefficient is analogous to the reliability coefficient in classical test theory [17].

Table 2. the g-coefficient for each manual method and the study of McClure et al. [13]

	Current Study		McClure's
	Reading	**Teaching**	
Structural Scoring	0.7520	0.9029	0.23
Relational Scoring w/o Criteria	0.8659	0.8540	0.51
Relational Scoring w/ Criteria	0.8874	0.9133	0.76

In this study, we interpret the g-coefficient as an estimate of score reliability assuming a single rater which shows the consistency of each scoring method as shown in Table 2. All values of G-coefficient of the current study are higher than values reported in McClure et al. [13]. Then, the scoring with the criteria map resulted in the highest score reliability in both reading and teaching situations, which is consistent with the investigation of McClure et al. which indicated that the relational scoring method is reliable in assessing the concept map. Based on these results, we concluded that the manual assessment conducted in this research is reliable and it is possible to evaluate reliability of KB map by comparing with the results of the manual assessment. As for the reason why the g-coefficient obtained in the current study is higher than that obtained by McClure et al. we guess that the current study was conducted with a smaller number of subjects and raters, that is, 12 raters in McClure et al., and 4 raters in the current study.

The reliability of KB method.
To confirm the reliability of the KB map, a comparison between the KB map's result and the reliable manual method's result is required. The Pearson's correlation was computed using the R programming language and the correlation value is shown in Table 3. Following the strength of the correlation from Evans [18], the relational method with criteria map, which achieved the highest reliability score, has a very strong correlation with the KB map in both reading and teaching situations. This is because raters use the criteria map as a frame for their scoring, in a similar way to the teacher-build map used in the KB map. For the remaining methods, the results from the relational scoring without criteria map have a very strong correlation in the reading situation and strong correlation in the teaching situation. This is because the procedure of relational scoring without the criteria map is too wide for meaningful evaluation of the learner-build maps, which are constructed for checking the understanding following a specific teaching situation. The structural scoring has a strong correlation with the KB map in both situations, even though structural scoring scores the concept map by giving precedence to the structure of the concept map, which is a different approach compared to the KB map.

The results above suggest that the KB map can assess learners' comprehension of a topic as well as the manual concept map assessment methods. If the manual methods give a relatively high score to a learner, the KB map also has a high possibility of giving a relatively high score to the learner. In addition, learners who get a relatively low score from the manual methods, also have a high possibility of getting a relatively low score from the KB map. As indicated by the high correlation value, the KB map

is reliable, and is comparable to the manual methods, in identifying learners' comprehension for a topic and evaluating the concept map.

Table 3. The correlations in scores between each manual method and the KB method

	KB in Reading	KB in Teaching
Structural Scoring	0.7360	0.7360
Relational Scoring w/o Criteria	0.8532	0.7371
Relational Scoring w/ Criteria	0.8671	0.8165

Note: Calculated Pearson product correlations are statistically significant as indicated by p-value < 0.01

4.2   Results of Questionnaire

Two sets of questionnaires were used in this study. The first questionnaire was for learners after they completed all of their tasks, this is presented in Table 4. From the learners' questionnaire analysis, learners who did not have existing knowledge about the learning material before obtained a good understanding of the content after reading. In addition, the learners could accept the instructor's interpretation clearly after they received an explanation in the teaching situation. When learners constructed their learner-build map by creating their own linking words, most of them concluded that they could represent their understanding adequately; similarly, users of the kit KB map were able to express their understanding appropriately. This summary suggests that the KB map is appropriate to use in supporting learners to express their understanding, and that it produces similar results to using the concept map where the linking words are created freely.

Table 4. A part of the learners' questionnaire

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Learners know about concept map before	9%	14%	9%	55%	14%
Learners know about material before	18%	27%	9%	41%	5%
Learners can represent their understanding by using CMapCloud	0%	5%	18%	73%	5%
Learners can represent their understanding by using KB map	0%	5%	0%	36%	59%

For the raters' questionnaire, all raters identified their familiarity with using the concept map and their understanding of the learning material as strong confident. In the raters' ranking of the manual methods, which is illustrated in Figure 8, the structural scoring was the hardest assessment method, because the rater had to decide on the suitability of each hierarchy and crosslink. Conversely, it was easiest to use the relational scoring with criteria map since the criteria map could be used as a guide for scoring. For the cost of scoring, the raters noted that the structural scoring and the relational scoring without criteria map used their memory load and time more than the

relational scoring with criteria map. This was because of the difficulty in thinking about the learner-build map structure and recalling how previous learner-build maps were scored. For this challenge, the criteria map can help the scoring of the learner-build map by using the relational scoring with the criteria map. In the final question, the raters were requested them to rank the most reasonable method in their opinion. The relational scoring with criteria map achieved the highest rating. This ranking corresponds with the comparison between six concept map assessments by McClure et al. [13]. Hence, the strong correspondence between the KB map and the relational scoring with criteria map confirms that the propositional level with exact matching of the KB map is a reliable method to assess the efficiency of learning and the KB map can be used as an alternative automatic method for assessing the concept map.
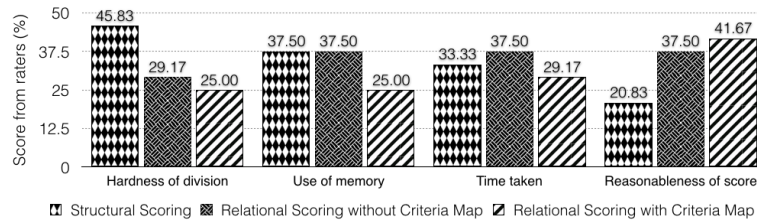

Fig. 8. A part of the score from the raters' questionnaire

5 Conclusion

This study investigates the reliability of the KB map in terms of its ability to identify the efficiency of learning. An experiment was designed to compare the KB map with three manual concept map assessment methods in reading and teaching situations. Selected manual methods contained structural scoring, (which investigates the composition of the concept map straightforwardly), relational scoring without the criteria map, and relational scoring with the criteria map. The relational scoring gives precedence to the meaning of propositions and is reasonable for evaluating understanding from a concept map, but requires expert checking and significant time input for scoring. These manual methods provide flexible and meaningful concept map assessment, and their reliability is widely accepted. However, they are inconvenient due to the limited class time that instructors have to complete a unit of instruction. In this study, the KB map was compared with the manual methods to test the assumption that the KB map is reliable in identifying the efficiency of learning. From this study, the results show a strong and significant correlation between the KB map and the manual methods in both the teaching and reading situations. The KB map has the highest correlation with the relational scoring with criteria map, achieving the most reliability score (g-coefficient) in both learning situations. Moreover, the learner-build map scores of the KB map were similar to the manual methods. Based on these results, it is concluded that the reliability of the KB map assessment is comparable to the manual methods.

References

1. Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. American Educational Research Journal, 28(1), 117-153.
2. Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). Educational psychology: A cognitive view (2nd ed.). New York: Holt, Rinehart and Winston.
3. Novak, J. D., & Cañas, A.J. (2008). Technical Report IHMC CmapTools. Florida: Institute for Human and Machine Cognition.
4. Hu, M. L. M., & Wu, M. H. (2012). The effect of concept mapping on students' cognitive load. World Transactions on Engineering and Technology Education, 10(2), 134-137.
5. Hirashima, T., Yamasaki, K., Fukuda, H., & Funaoi, H. (2011). Kit-Build Concept Map for Automatic Diagnosis. Proceedings of Artificial Intelligence in Education 2011 (pp.466-468). Auckland, New Zealand: Springer-Verlag Berlin Heidelberg.
6. Hirashima, T., Yamasaki, K., Fukuda, H., & Funaoi, H. (2015). Framework of Kit-Build concept map for automatic diagnosis and its preliminary use. Research and Practice in Technology Enhanced Learning, 10(1), 1-21.
7. Sugihara, K., Osada, T., Nakata, S., Funaoi, H. & Hirashima, T. (2012). Experimental evaluation of Kit-Build concept map for science classes in an elementary school. Proceedings of Computers in Education 2012 (pp.17-24). Singapore: National Institute of Education.
8. Yoshida, K., Sugihara, K., Nino, Y., Shida, M., & Hirashima, T. (2013). Practical Use of Kit-Build Concept Map System for Formative Assessment of Learners' Comprehension in a Lecture. Proceedings of Computers in Education 2013 (pp.906-915). Bali, Indonesia: Asia-Pacific Society for Computers in Education.
9. Nomura, T., Hayashi, Y., Suzuki, T. & Hirashima, T., (2014). Knowledge Propagation in Practical Use of Kit-Build Concept Map System in Classroom Group Work for Knowledge Sharing. Proceeding of International Conference on Computers in Education Workshop 2014 (pp.463-472). Nara, Japan: ICCE 2014 Organizing Committee.
10. Alkhateeb, M., Hayashi, Y. & Hirashima, T. (2015). Comparison between Kit-Build and Scratch-Build Concept Mapping Methods in Supporting EFL Reading Comprehension. The Journal of Information and Systems in Education, 14(1), 13-27.
11. Novak, J. D., & Gowin, D.B. (1984). Learning how to learn, New York: Cambridge University Press.
12. McClure, J.R., & Bell, P.E. (1990). Effects of an environmental education related STS approach instruction on cognitive structures of pre-service science teachers. Pennsylvania: State University.
13. McClure, J.R., Sonak, B. & Suen, H.K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. Journal of Research in Science Teaching, 36(4), 475–492.
14. Klaus, R. (2013). Sugar. Retrieved on August 01, 2016, from http://www.english-online.at/biology/sugar/sugar-carbohy-drate-that-gives-us-energy.htm
15. The Institute for Human & Machine Cognition. (2014). Cmap Cloud & CmapTools in the Cloud. Retrieved on August 2016, from http://cmap.ihmc.us/cmap-cloud/
16. Brennan, R. L. (1983). Elements of generalizability theory. American College Testing Program.
17. Webb, N. M., & Shavelson, R. J. (2005). Generalizability theory: overview. Wiley StatsRef: Statistics Reference Online.
18. Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. Brooks/Cole.