Rismanto et al. *Research and Practice in Technology Enhanced Learning* (2024) 19:21
https://doi.org/10.58459/rptel.2024.19021

Research and Practice in Technology Enhanced Learning

RESEARCH Free and Open Access

Evaluating the kit-build concept mapping process using sub-map scoring

Ridwan Rismanto [1,2*], Aryo Pinandito [3], Banni Satria Andoko [2], Yusuke Hayashi [1] and Tsukasa Hirashima [1]

*Correspondence:
rismanto@polinema.ac.id
[1] Informatics and Data Science, Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1 Kagamiyama, 739-8527, Higashihiroshima, Japan
[2] Department of Information Technology, State Polytechnic of Malang, Jl. Soekarno Hatta No. 9, Lowokwaru, 65141, Malang, Indonesia
Full list of author information is available at the end of the article

Abstract

Concept mapping allows learners to visually represent their knowledge by connecting nodes (concepts) and links (relations between concepts). The kit-build (KB) concept map framework enhances this process by enabling learners to recompose a concept map from provided nodes and links, leading to improved learning outcomes. Additionally, KB employs an automatic assessment method called "Full Map Scoring (FMS)", which evaluates the learner's understanding based on the recomposed concept map. However, FMS only evaluates the final product of the recomposition activity, neglecting the process itself. This is a potential limitation because different processes leading to the same result could reflect different levels of understanding among learners. Therefore, it is crucial to incorporate process analysis into learner assessment. To address this issue, our research proposes a new assessment procedure termed "Sub-Map Scoring (SMS)". A concept map is generally composed of several sub-maps with each sub-map representing a set of meanings. We hypothesize that if a learner comprehends the meaning of a sub-map, the learner will recompose the sub-map as a continuous activity. Therefore, SMS evaluates the recomposition process of each sub-map from the viewpoint of continuity, and the overall SMS score is derived from these evaluations. To verify the effectiveness of SMS, we compared SMS and FMS scores using data from a practical use of the KB framework. A multiple linear regression analysis confirmed that the SMS score was a more precise predictor of learning gain than the FMS score.

Keywords: Concept map, Kit-build concept map, KB activity process, Sub-map

Introduction

A concept map is a well-known visual thinking technique popularly used in education (Jonassen et al., 2013; Novak, 1990; Novak et al., 1984). In an educational context, it is popularly used to request learners to visually represent their understanding of an object of learning (Cañas et al., 2016; Dwi Prasetya et al., 2020). Many investigations have reported that the activity of representing understanding with the concept map enhances learning more effectively than traditional learning methods (Elhelou, 1997; Horton et al., 1993; Lee et al., 2013).

Moreover, various studies have been conducted using the concept map as an assessment tool to evaluate learner's understanding (Gregoriades et al., 2009; Kinchin et al., 2000; Vanides et al., 2005). Most of these studies have relied on manual assessment methods conducted by human raters and have reported their usefulness. However, manual assessment methods present two issues, they are time-consuming and lack the stability of results among raters (Papajohn, 2002; Ruiz-Primo & Shavelson, 1996). As a solution of these problems, automatic assessment has been investigated as a promising approach (Wunnasri et al., 2018b).

Kit-build concept map (KB) is a framework that realizes automatic assessment of concept maps made by learners (Hirashima et al., 2015; Yamasaki et al., 2010). In KB, learners are provided with a set of components generated by decomposing a concept map created by a teacher, which is made as an ideal representation of the understanding learners should achieve. The learners are then requested to recompose a concept map by connecting the provided components. The map recomposed by the learner is automatically assessed by comparing it with the original teacher's map.

The validity of the automatic assessment by KB has been confirmed through comparisons with manual evaluation methods (Wunnasri et al., 2018b). Additionally, the usefulness of the diagnosis for additional teaching by teachers (Pailai et al., 2017; Yoshida et al., 2013), automatic feedback by learning support systems (Furtado et al., 2019; Pailai et al., 2018), and pair discussion (Sadita et al., 2020; Wunnasri et al., 2018a) have been investigated through experiments and practice.

However, the assessment of KB is specifically conducted only for the result of the recomposition, and the process of the recomposition itself is not the target of the assessment. Therefore, if learners recompose the same map, but with different recomposition processes, they will receive the same assessment result. Several previous investigations, however, have highlighted that in the context of a learning activity, similar outcome can be led to a different level of understanding when the process is different (Ford et al., 1998; Lee & Fortune, 2013; Van Rossum & Schenk, 1984). As noted by Marton and Säljö (1976), process analysis could reveal relations to the degree of understanding and learning gain. Especially, because of the nature of KB that enables learners to recompose the same map in various ways, the process assessment is an important research issue to adequately evaluate learner's understanding.

In this study, we focus on the idea that a concept map can be interpreted as a combination of several sub-maps, each representing a set of meanings (Kinchin & Alias, 2005; Roberts, 1999). Based on this idea, it is reasonable to assume that learners will recompose the sub-map through a series of activities if they have a good understanding of the set of meanings. Based on this assumption, it becomes possible to evaluate the process of sub-map recomposition. If this assumption and approximate evaluation of sub-map are adequate, it becomes possible to assess the process of the whole map recomposition as the sum of the assessments of each sub-map recomposition process. The assessment method used in previous research, targeting the whole concept map as an outcome of recomposition, is called Full Map Scoring (FMS for short). The new method proposed in this paper, focusing on recomposition processes of sub-maps is called Sub-Map Scoring (SMS).

SMS has been implemented and evaluated in comparison with FMS. The evaluation was conducted using data obtained from the practical use of KB in a university class with 60 second-year undergraduate students. The learning gain as the normalized differences between pre-test scores and post-test scores was used in this evaluation. Hence the first research question (RQ1) is: Can SMS assess the process of KB concept mapping activity in terms of its relationship with the learning gain? The second research question (RQ2) is: Is SMS a better predictor of learner's learning gain compared to FMS?

This paper's Background subsections explain the following: (a) the basics of a concept map as an assessment tool, (b) the KB concept map as the framework for concept map recomposition and automatic diagnosis and (c) sub-map in a concept map. The Sub-Map Scoring section discusses the approach used in assessing the KB concept map recomposition process. The Experimental Evaluation section describes the participants, instruments, procedures, and analysis of the experiment. The Results and Discussion section centers on the experiment results and deliberation of our findings. Lastly, the Conclusions and Future Works section summarizes this research and discusses future works.

Background

Concept map as an assessment tool

Concept maps are represented by graphs and consist of two symbols: nodes represent concepts and links represent the relationship between concepts (Cañas et al., 2016; Dwi Prasetya et al., 2020). In the context of learning subjects and reading materials, relationship between concepts represents proposition. In educational context, concept map is a widely used technique for learning and assessment (Hirashima et al., 2015; Plotnick, 1997; Vanides et al., 2005; Zheng et al., 2019). Several investigations have mentioned the advantages of concept map, including the ability to enable knowledge sharing (Novak & Symington, 1982), improve learner's creativity and motivation (Chan, 2017), improve critical thinking skills (Sundararajan et al., 2018), and actively engage learners in the learning process (Novak et al., 1984).

As an assessment tool, many methods have been proposed to utilize concept maps to assess learner's knowledge. Several research have confirmed the validity and reliability of using concept map as an assessment tool (Plotnick, 1997; Vanides et al., 2005; Zheng et al., 2019). There are two main approaches in assessing the concept map: one is structural scoring and the other is relational scoring. A structural scoring was proposed by Novak et al. (1984). This method works by assessing the correctness of hierarchy level and crosslink validity which reflects learner's ability in creative thinking (West et al., 2000). Relational scoring, on the other hand, considers the meaning of propositions instead of the map structure (McClure & Bell, 1990). This method works by evaluating the possibility of relationship between each proposition, appropriateness of concept labels and compatibility between labels.

McClure et al. (1999) conducted a study between the two methods by requesting 63 students to construct a concept map using 20 provided concepts and creating the links to connect the concepts. A group consisting of 12 raters then scored each map separately. Result shows that the relational and structural scoring method has a close relation when a criteria map, or teacher-build map is used as a scoring reference. Kit-build concept map also uses the expert map as a criteria map for assessment (Hirashima et al., 2015). The automatic assessment in the KB map is explained in the following subsection in more detail.

Kit-build concept map

The kit-build concept map is a concept-mapping framework that enables the automatic diagnosis of a learner's concept map (Hirashima et al., 2015) that consists of teacher's map building, learner map building, and KB analyzer. In this framework, there is an activity in which the learner is tasked with recomposing the concept map components (kit), which consists of nodes and links, into a complete map. The kit was prepared by decomposing the teacher's map, which is kind of expert map created by a teacher for the learning material. In KB concept map, learners recompose the teacher's map instead of creating their own concept maps from scratch.

In the kit-build concept map, the goal-map building is divided into two tasks: segmentation and structuring (Yamasaki et al., 2010). The segmentation task involves extracting the nodes and links (kit) of a concept map from the learning materials, while the structuring task involves reconnecting the extracted elements (kit) and creating a complete concept map. Figure 1 illustrates how the kit (i.e., concepts and links) was extracted from the learning materials or resources during the segmentation task and how the kit's elements were connected to form the concept map during the structuring task.

The next step following the segmentation and structuring task, the goal map is decomposed into a kit of separated nodes and links, which is then given to a learner. The learner recomposes the nodes and links into a map which is called "learner map"

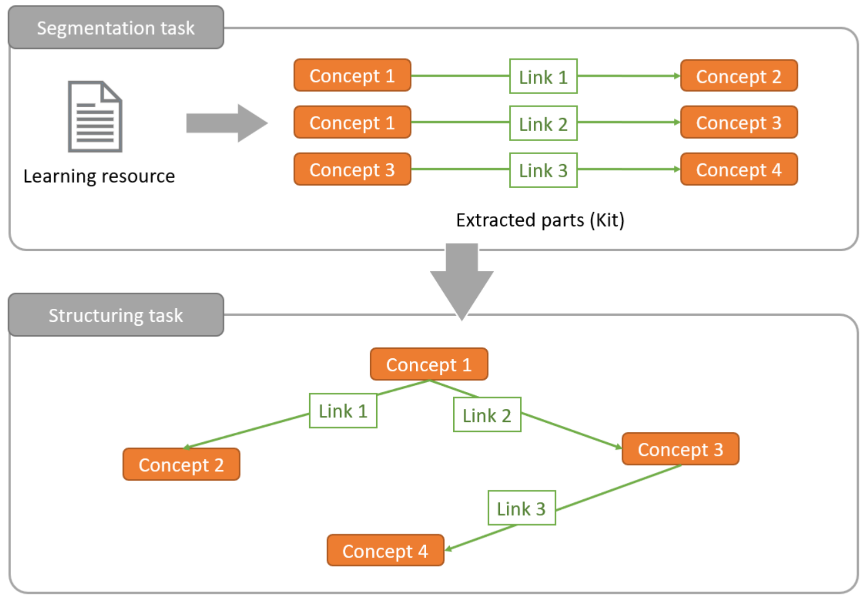

Fig. 1 Segmentation and structuring tasks

(structuring task). After the learners complete the map recomposition, the system automatically compares the learner's map to the goal map in terms of their similarities. The diagnosis works by superimposing the learner map over the goal map to detect which and how many links are excessive (i.e., incorrectly connected), missing (i.e., should be connected but not), and matching (i.e., correctly connected) (Hirashima et al., 2015). Figure 2 shows the kit-build concept map system's lifecycle flow. In the Phase 3 of the Figure 2 illustrates the diagnosis. Blue-solid line means excessive links, red-dashed line means missing links, and green line means matching links.

The framework also allows the automatic diagnosis of several learner maps, known as group maps (Hirashima et al., 2015). In this diagnosis, multiple learner maps are combined into a group map, which is then overlaid with the goal map. By leveraging this kind of diagnosis, teachers can determine which parts of the learning material are not properly understood, misunderstood, and accurately understood by the majority of the learners and receive apt feedback to assess the learning condition.

KB has advantages over scratch-build (SB) concept map as suggested by several research (e.g., Alkhateeb et al., 2015; Andoko et al., 2020; Funaoi et al., 2011). Compared to SB, components to build the concept map are provided in the KB system and it allows analysis on the recorded learner's activity, including the connected propositions in relation to the teacher's map. Therefore, the proposed method in this study was applied to the KB concept mapping.

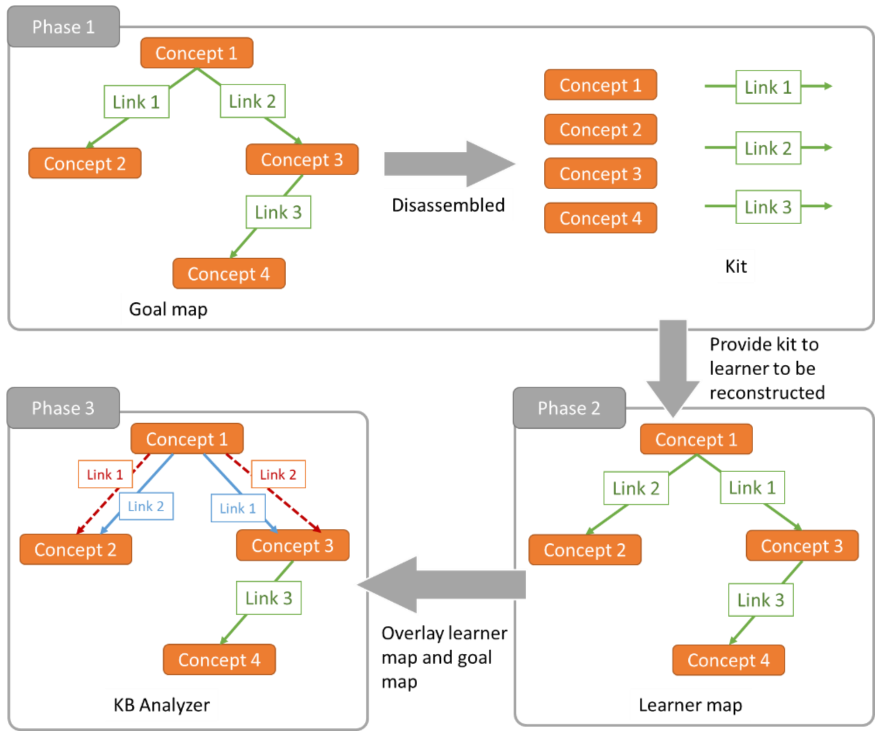

Fig. 2 The kit-build concept map system's lifecycle

Sub-map in a concept map

In a concept map, propositions that share similar meanings and belong to a specific domain of knowledge can be grouped together into a sub-map. A sub-map typically comprises propositions that have related meanings are derived from the same subdomain of knowledge. According to Gerstner and Bogner (2009), sub-maps are defined as groups of propositions originating from distinct domains within the learning material. Schneider et al. (2021) states that dividing a concept map into a series of sub-maps can be accomplished by focusing on a specific knowledge domain. This approach aligns with the spatial continuity principle described by Moreno and Mayer (1999) and Schroeder and Cenkci (2018), which suggests arranging propositions that share similar information in close proximity to one another within the concept map.

Additionally, segmenting the learning material into meaningful and coherent segments, as recommended in the segmenting principle, can enhance learning outcomes (Schneider et al., 2021). In a study conducted by Kinchin and Alias (2005), a sub-map is described as a collection of propositions with meaningfully related concepts. The study highlights that a teacher created the sub-map to provide a concise overview of the topic's content. Defining the sub-map can also aid in identifying prerequisite knowledge necessary to understand a particular sub-map effectively.

Sub-Map Scoring (SMS)

The kit-build concept map has already incorporated the Full Map Scoring (FMS) to score the concept maps recomposed by learners. The map recomposed by a learner is called a learner map. This scoring method compares propositions in a learner map to those in the goal map. Because both maps are composed of the same components, if a proposition exists in both maps, the proposition is judged as a correct proposition. If a proposition exists only in the learner map, it is judged as an excess proposition. A proposition existing only in the goal map is judged as a lacking proposition. The score would be calculated by counting the number of correct propositions and comparing it to the number of propositions from the teacher's map. For example, when the number of correct propositions is five and the number of the teacher's map's propositions is ten, the FMS score is 50.

In contrast to FMS, SMS has been proposed based on the idea that the structure of a concept map consists of several sub-maps, a sub-map is a subset of propositions with tightly related meanings. The proposed method investigates the learner's learning gain based on recomposition process of each sub-map. As a preparation of SMS, a teacher is required to define several sub-maps in the teacher's map. For each sub-map, SMS evaluates the learner's recomposition sequence of propositions belonging to the sub-map. Figure 3 illustrates the sub-map specified by the teacher in the goal map, while Figures 4 and 5 depict the sequence of propositions recomposed by learners.

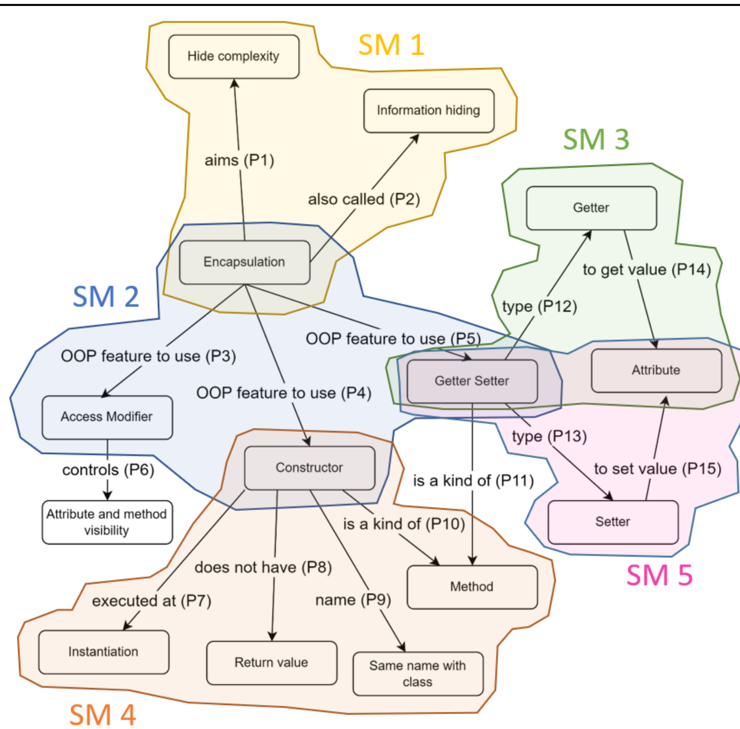

Fig. 3 Sub-map specified within a teacher's map

Figure 3 shows five sub-maps (SMs) defined by a teacher. Sub-map 1 (SM1) is made up of two propositions: proposition 1 (P1) and proposition 2 (P2), which are written in the parentheses at the end of a link label. SM2 comprises P3, P4, and P5, while SM3 involves P12 and P14. SM4 has P7, P8, P9, and P10, while SM5 covers P13 and P15. This sub-map configuration can be written as follows: {P1, P2}{P3, P4, P5}{P12, P14}{P7, P8, P9, P10}{P13, P15}.

Figures 4 and 5 illustrate two examples of different proposition recomposition sequences. The order of proposition recomposition is represented by numbers in rectangles, such as O1 or O2. O1 means the first connected proposition, O2 means the second connected proposition, and so on. For the sake of clarity, only the propositions used from O1 to O7 are shown. Figure 4 displays the work of two learners (Learners A and B) who recomposed the same final map, with their recomposition processes differing in the proposition sequence order from O1 to O7.

Since the two recomposed maps in Figure 4 are identical, both recomposed maps get the same FMS score although their recomposition processes are different. SMS is able to give difference scores for the two recompositions. By using these examples, the procedure is explained concretely. It is assumed that Learner A completed each of the three sub-maps in continuity, as shown in Figure 5 (top). In the figure, Learner A connected P1 and P2, which belong to sub-map 1 (SM1). Consecutively, P3, P5, P4, which belong to SM2, then P14, P12 which belong to SM3. The manner in which Learner A connected propositions that belong to a sub-map consecutively indicates a good understanding of the set of propositions as a unit of knowledge. When a learner organizes and presents the propositions in a coherent and consecutive manner, it suggests that they grasp the relationships and connections between those propositions.

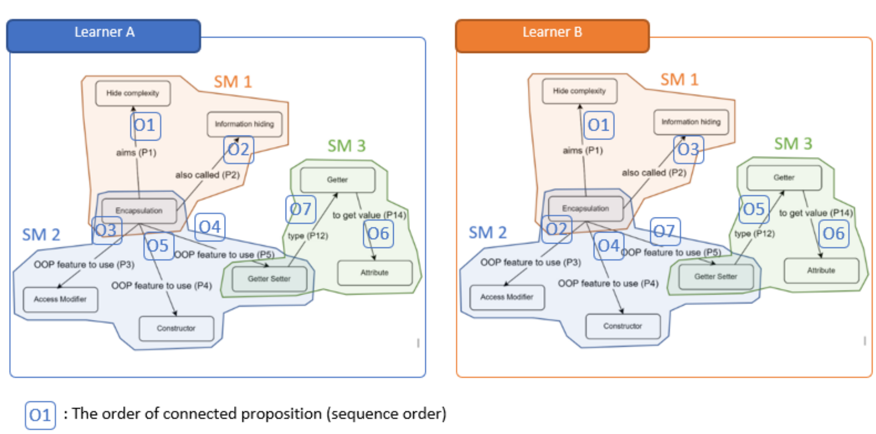

Fig. 4 Sequence example for Learner A (left) and Learner B (right)

In contrast, Learner B completed the sub-maps in discontinuity, as shown in Figure 5 (bottom). In the figure, Learner B connected P1, P3, P2, P4, P12, P14, P5, in which only P12 and P14 were connected consecutively that belong to a sub-map (SM3). These processes imply that Learner B comprehends the set of propositions as a unit of knowledge less than Learner A. Therefore, Learner A's process score should be higher than that of Learner B.

In the context of formative assessment, teachers could use the process assessment to evaluate learner's process in the activity of recomposing the concept map. For example, the teacher may encourage Learner B to further enhance their understanding of the learning material. The discrepancy in connected propositions suggests that Learner B might not be effectively elaborating on the knowledge implied in a set of propositions that share similar meanings.

Based on the above consideration, we proposed Sub-Map Scoring (SMS) as the recomposition process's scoring method. To calculate the SMS, we analyzed the propositions' sequence connected by the learner. In each sub-map, we scanned the learner's proposition that is sequentially connected and matched the sub-map's propositions. Figure 5 illustrates an example of the analysis. In the teacher's map, there are three

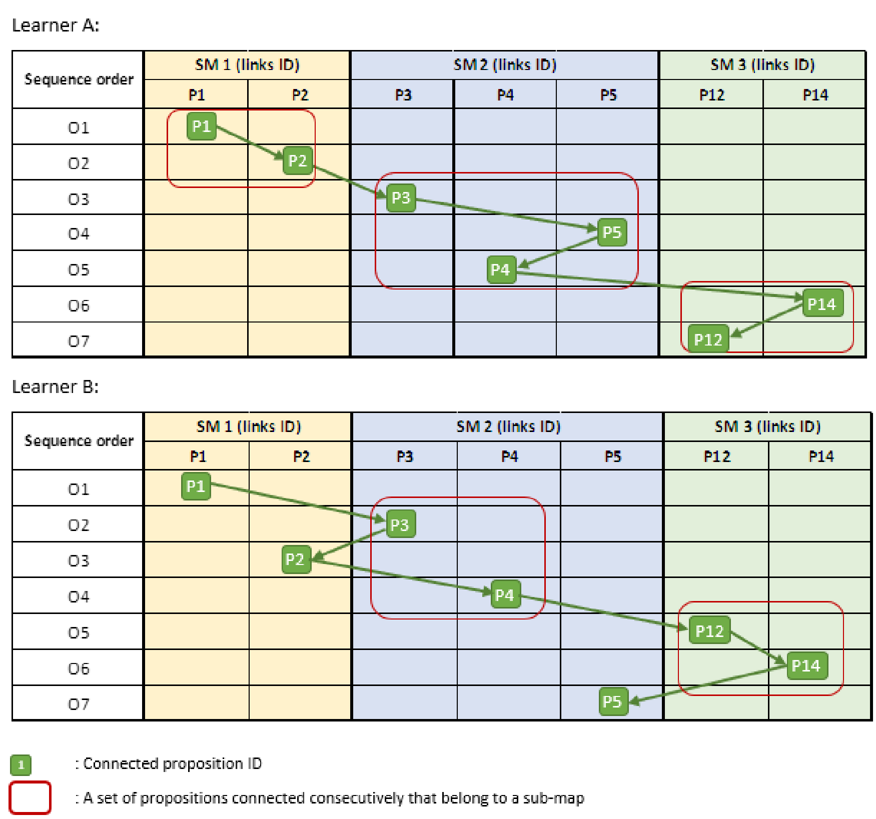

Fig. 5 Sequence path summary example for Process A and Process B

sub-maps: SM1: {P1, P2}, SM2: {P3, P4, P5}, and SM3: {P12, P14}. The Learner A recomposition sequence is {P1, P2, P3, P5, P4, P14, P12}.

According to this data, the matched proposition sequence with the sub-maps is {P1, P2} for SM1, {P3, P5, P4} for SM2, and {P14, P12} for SM3. The method treats the propositions in a sub-map as a set, therefore, the percentage of matched propositions sequence for each sub-map is then calculated. If a sub-map is composed of four propositions, two of which are recomposed in continuity, the sub-map's score is calculated as (2 / 4) * 100 = 50. The proposed method assumes that all the sub-maps have equal weight of importance, so the total SMS is the average of each sub-map score. This percentage is the rate of the number of propositions connected as a set, in comparison with the number of propositions in a sub-map. This is to represent the rate of sub-map completeness. The complete calculation for the two learners is as follows:

Learner A

SMS Score: ((2 / 2 * 100) + (3 / 3* 100) + (2 / 2 * 100)) / 3 = 100.

Learner B

SMS Score: (0 + 0 + (2 / 2 * 100)) / 3 = 33.

In the case of a sequence with recurring propositions set that belongs to a sub-map, the selection is based on the set with the greatest number of propositions. This is because a larger set of connected propositions is assumed to better reflect an understanding of the knowledge context conveyed in the sub-map. For example, a sub-map 1 (SM1) consists of {P1, P2, P3, P4, P5} and a learner made a sequence of {P1, P2, P9, P10, P3, P4, P5}. In this sequence, {P1, P2} matched the SM1, and the next sequence is {P9, P10}, is no longer matched SM1. However, the next sequence is {P3, P4, P5}, which recurringly matched the SM1. In this case, the sequence of {P3, P4, P5} is counted to be scored.

Experimental evaluation

Participants

The subjects of this research are 60 second-year undergraduate learners from the Information Technology Department of the State Polytechnic of Malang, Indonesia. The students have no prior knowledge or experience building a concept map. Those who did not complete the experiment activities, did not attend the experiment sessions due to a system error, or left the question tests blank were not included. After the teacher delivered the learning material presentation, the learners used the kit-build system.

Instruments

The experiment's instruments included the learning module used in the Object-Oriented Programming (OOP) course. This module was compiled and released by the OOP course teaching team (four teachers) from the Information Technology Department of the State Polytechnic of Malang, Indonesia. One of the team members is the first author. The module material complied with the standard OOP education provided by the ORACLE academy program. The OOP concepts used were extracted from the module's encapsulation topic.

The kit-build goal map used for this experiment consists of 15 propositions with 14 concepts, 15 links, and five sub-maps, as seen in Figure 3. The goal map was proposed by this paper's first author and presented to the three OOP teaching team members. The goal map was then discussed and modified according to the teaching team's inputs. The pretest, posttest, and delayed test employed the same 20 multiple-choice questions, all of which were collaboratively prepared by the teaching team.

Procedures

The experiments took a total of 60 minutes. The learners were first given an introduction and tutorial on how to use the KB system. At first, learners were given a pretest to measure their learning knowledge baseline. The pretest was completed in 10 minutes. Afterward, the learning material was delivered in the form of a 15-minute teacher presentation. Then, the kit-build concept mapping activity using the KB system followed.

In the KB system, learners were tasked to reconstruct the kit or concept map components (concepts and links) into a learner map, which took 15 minutes to complete. A 10-minute posttest then followed. As mentioned, the questions on this test are the same as those on the pre-test. A week later, using the same questions, the delayed test was administered in 10 minutes. During the pretest, posttest, and delayed test, all the materials were closed.

Analysis

In this research, six variables were used to perform the analysis: pre-test score, post-test score, delayed-test score, learning gain, SMS score, and FMS score. The learning gain was calculated and used to evaluate the learner's understanding by measuring it based on pre-test score and post-test score using normalized gain formula, introduced by Hake (1998). The normalized gain measurement has some significant advantages over the regular gain calculation in which gain = post-test score – pre-test score (Bao, 2006; Coletta & Steinert, 2020; Marx & Cummings, 2007). Particularly, it allows teachers to compare increases in scores for many learner populations by taking the pre-test score into consideration and thereby focusing on the possible improvement which makes learners with different pre-test scores more comparable. The normalized gain formula is as follows:

$$NLG = (PostScore - PreScore)/(100 - PreScore)$$

The learning effect of the KB concept mapping activity was evaluated by analyzing significance of differences between pre-post-delayed test score using analysis of variance (ANOVA). Shapiro-Wilk test was performed on the test results and resulted in a normal distribution on the pre-test scores (p-value = 0.2705), non-normally distributed data on the post-test scores (p-value = 6.774e-06) and delayed-test scores (p-value = 0.006). Therefore, a non-parametric ANOVA using Kruskal-Wallis test was performed to evaluate the KB activity effect. A p-value of < 0.05 considered statistically significant differences between scores.

The correlation test was performed to evaluate the relationship between SMS, FMS, post-test score, delayed-test score and learning gain using Pearson's r coefficient. To interpret the r-value, we use Schober et al. (2018) criteria of $r < 0.39$ being weak, $r > 0.40$ and $r < 0.69$ being moderate, and $r > 0.70$ being strong correlation. Multiple linear regression analysis was performed to evaluate between SMS and FMS in terms of predicting learner's learning gain.

Results and discussion

Table 1 summarizes the experimental data, showing the mean and standard deviation values of the pre-test score, post-test score, delayed-test score, learning gain, SMS, and FMS.

Kit-build activity effect analysis

Before evaluating the SMS, we performed the KB activity effect analysis to confirm that the activity produced adequate outcome that was observed in previous research (e.g., Andoko et al., 2020; Dwi Prasetya et al., 2020; Funaoi et al., 2011; Hirashima et al., 2015; Pailai et al., 2017; Yamasaki et al., 2010). This analysis compared the pre–post test scores and the pre–delayed test scores to investigate the KB concept mapping's activity effect. Figure 6 shows the test score visualization of the pre-test, post-test, and delayed-test scores. The mean analysis results are depicted in Table 1.

The mean analysis results in Table 1 show that the pre-test mean score was 33.960, the post-test mean score was 78.118, and the delayed-test mean score was 73.633. To assess

Table 1 Mean and standard deviation of the pre-test score, post-test score, delayed-test score, learning gain value, SMS, and FMS

N	Pre-test (0 – 100)		Post-test (0 – 100)		Delayed-test (0 – 100)		Learning gain (-1 – 1)		SMS (0 – 100)		FMS (0 – 100)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
60	33.96	13.33	78.12	19.87	73.63	19.26	0.68	0.28	62.98	26.99	73.95	28.90

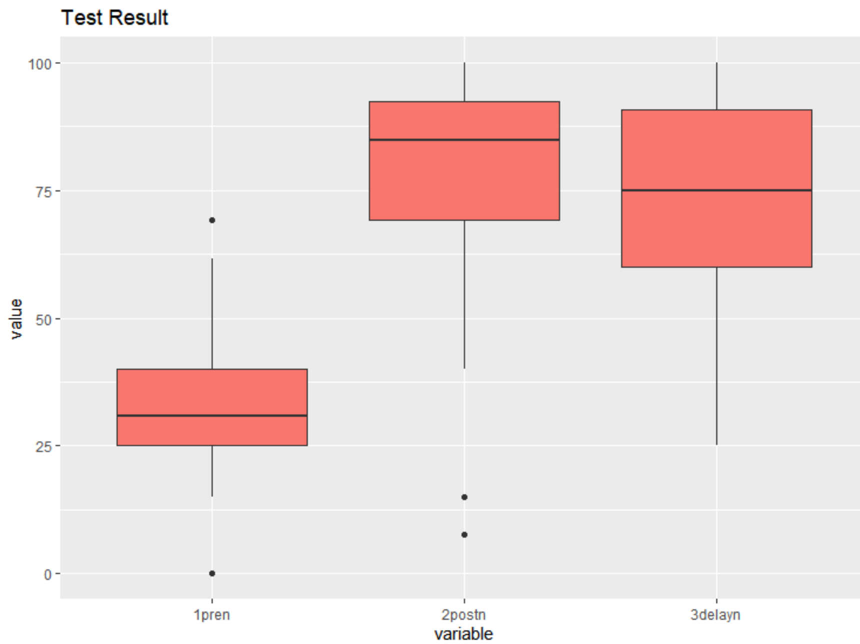

Fig. 6 Pre-test, post-test, and delayed-test scores

the significance of the pre-test and post-test mean scores, as well as the post-test and delayed-test mean scores, a nonparametric ANOVA using Kruskal–Wallis analysis was conducted, and the results are displayed in Table 2. With a p-value < 0.001, the post-test score is statistically significant than the pre-test score. Post-test and delayed-test scores were not statistically significant, with a p-value = 0.194, $p > 0.05$, as shown in Table 2.

Correlation analysis

To answer the first research question (RQ1), correlation test of learning gain with the SMS and FMS was used to evaluate the relationship between the scores with learning gain. The results are shown in Table 3. The correlation value between the SMS and learning gain was 0.711, indicating a strong positive relationship, while the value between the FMS and learning gain was 0.674, indicating a moderately positive link. Both are statistically significant with a p-value < 0.05.

Table 2 Kruskal-Wallis analysis results

	chi-squared	*df*	*p*-value
Pre–Post test	72.022	1	< 0.001
Pre–Delay test	68.484	1	< 0.001
Post–Delay test	1.6872	1	0.194

Table 3 Correlation between the SMS and learning gain, and the FMS and learning gain

	t	*df*	*p*-value	*r* coefficient
SMS – Learning Gain	7.7051	58	1.91E-10	0.711218
FMS – Learning Gain	6.9489	58	3.57E-09	0.674025

Multiple linear regression analysis

To answer the second research question (RQ2) in discerning how the scores (i.e., SMS and FMS) predict the learner's learning gain in terms of the learning material delivered by the teacher, a multiple linear regression analysis (MLRA) was performed.

First, the pre-test score was examined to determine whether it also predicts learning gain. Regression analysis was performed on the pre-test score and FMS, as well as on the pre-test score and SMS. The first test of pre-test score and FMS to predict learning gain, results showed that pre-test has a *p*-value of 0.392, FMS has a *p*-value of less than 0.05 with $R^2 = 0.461$. In the second test of pre-test score and SMS to predict learning gain, results showed that pre-test has a *p*-value of 0.173, SMS has a *p*-value of less than 0.05 with $R^2 = 0.521$. This indicates that when compared to FMS and SMS, the pre-test score cannot be used to predict the learning gain.

Second, the pre-test score, SMS, and FMS were compared to discern which among the three variables is the better predictor on the learning gain value. Table 4 shows that SMS, with a *p*-value of 0.00959, $p < 0.05$, and $R^2 = 0.5226718$, has a significant association with learning gain when compared to the pre-test score and FMS.

The Durbin-Watson test was performed to check the assumption that the residuals from the MLRA model are independent. The test resulted in *d*¬-value of 1.842 indicates that there is no autocorrelation problem. The *p*-value of 0.386 indicates that the residuals are independent and therefore the predictor variables are statistically significant. A further test on homoscedasticity using Non-constant Variance Score Test was performed with resulted in *p*-value of 0.148. This indicates that the residuals are homoscedastic, therefore the residuals have equal variance for every value of the fitted values and of the predictors.

Table 4 Multiple linear regression analysis on the pre-test score, SMS, and FMS in relation to learning gain

	Estimate	Std. error	*t*	*Pr*(>|*t*|)
(Intercept)	12.69902	9.13224	1.391	0.16986
Pre-test score	0.25228	0.19589	1.288	0.20308
SMS	0.65289	0.2434	2.682	0.00959**
FMS	0.07325	0.22863	0.32	0.74987

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Supplementary analysis

Additional analysis to evaluate the relationship between SMS, post-test score, delayed-test score and FMS was performed to confirm the usability of SMS in terms of assessing the learning activity outcomes and map quality.

To evaluate the relationship between learner's understanding in terms of the post-test score and SMS, and the FMS and post-test score, a correlation test was performed. The results showed that the SMS and FMS have a similar correlation value, as seen in Table 5, with r coefficients of 0.719 and 0.707, respectively, both of which indicate a strong positive correlation and are statistically significant with a p-value < 0.05.

The delayed-test score was used to measure the learner's memory retention after one week of receiving the learning material. The correlation between the SMS and delayed-test score, and the FMS and delayed-test score were also assessed to evaluate how SMS and FMS relate to memory retention. The results are depicted in Table 6. The correlation value of the SMS and delayed-test score was 0.588, which was moderately positive; similarly, the correlation value of the FMS and delayed-test score was 0.550, which was also moderately positive. Both have a statistically significant correlation with a p-value < 0.05.

A correlation test between SMS and FMS was conducted to confirm the usability of SMS in relation to FMS and to evaluate that learners who have a good process on the KB concept mapping activity means they will result in a good concept map quality. Table 7 depicts the relationship between SMS and FMS. The two variables are highly correlated with a correlation value of 0.919. The correlation is also statistically significant with a p-value < 0.05.

Table 5 Correlation between the SMS and post-test score, and the FMS and post-test score

	t	*df*	*p*-value	*r* coefficient
SMS – Post-test	7.8775	58	9.83E-11	0.718949
FMS – Post-test	7.6189	58	2.67E-10	0.707253

Table 6 Correlation between the SMS and delayed-test score, and the FMS and delayed-test score

	t	*df*	*p*-value	*r* coefficient
SMS – Delayed-test	5.5374	58	7.78E-07	0.588079
FMS – Delayed-test	5.0236	58	5.16E-06	0.550627

Table 7 Correlation between the SMS and FMS

	t	*df*	*p*-value	*r* coefficient
SMS – FMS	17.734	58	< 2.2e-16	0.918858

Discussion

Research question 1 & 2

As shown in the Kit-Build Activity Effect Analysis subsection, the results from the learning activity effect analysis confirmed that the KB activity outcomes were in-line with previous research, with a significant increase between pre-test and post-test scores. Therefore, the activity resulted in a meaningful outcome, hence it is adequate to proceed in analyzing the SMS based on these outcomes.

SMS, as well as the FMS, have a high correlation with learning gain with correlation value of 0.711 and 0.674 respectively. The FMS result is in-line with previous research that confirmed the relation between KB map score method with learner's understanding (Hirashima et al., 2015; Wunnasri et al., 2018b; Yamasaki et al., 2010). Based on this, it is implied that SMS and FMS can both be used as scoring systems to assess learner's learning gain.

Furthermore, multiple linear regression analysis shows a *p*-value of 0.00959 for SMS, when tested against FMS and pre-test score. This means that SMS is a better predictor of learning gain than FMS and pre-test score. Because learning gain is closely connected to the process of KB concept mapping activity, this suggests that the SMS is effective in evaluating the process in KB concept mapping.

A correlation value of 0.719 between post-test and SMS and 0.707 between post-test and FMS means that both SMS and FMS have a similar relationship in reflecting learner understanding represented by the post-test score. In terms of memory retention, a correlation value of 0.588 between delayed-test and SMS and 0.550 between delayed-test and FMS indicates that the SMS and FMS have the same relationship in reflecting the learners' memory retention. The similar relationship of SMS and FMS towards post and delayed test score indicates that learner's process in KB concept mapping affects the understanding and memory retention as well as the learner's map quality. Furthermore, the result of correlation analysis between FMS and SMS confirmed that learner's process in concept map building has a relationship with map quality (Chiu & Lin, 2012; Srivastava et al., 2021). Based on these results, it is implied that a good process in KB concept mapping could lead to not only a better learning gain, but also a better understanding and memory retention of the learning subject, and better map quality. Therefore, supporting the process in KB concept map recomposition could produce an adequate outcome in the activity.

Formative assessment based on process assessment

In the context of formative assessment, teachers could use the SMS as an assessment to evaluate the learner's knowledge elaboration. Knowledge elaboration has been defined as the interconnection, restructuring and integration of new information with prior knowledge

(Weinstein & Mayer, 1986; Zheng et al., 2022). In comparison with learner's understanding which reflects comprehension and interpretation of an information, knowledge elaboration refers to the process of building on existing knowledge to deepen one's understanding, make connections between different pieces of information, and apply knowledge in new and more complex situations. Good process of knowledge elaboration allows coherent structure of knowledge organization and are essential for meaningful learning (Kalyuga, 2009).

In a concept mapping activity, a process of knowledge elaboration can be evaluated in the process of concept map recomposing activity. This can be achieved by assessing how learners connect propositions as a process of organizing, restructuring and interconnecting concepts using their knowledge. In a study about knowledge acquisition and recall, Bergman et al. (2015) mentioned that the learner's ability to create a meaningful connection between propositions in a relevant context shows a good elaboration of learner's knowledge. In SMS, propositions that have a relevant context are grouped into a sub-map, and learners are scored based on the continuity of the connected propositions in a sub-map. Therefore, high SMS means that the learners are able to create a meaningful connection between propositions in a relevant context, hence have a good process of knowledge elaboration.

Moreover, even though the scoring method of SMS is different with FMS, the SMS could result in a similar correlation between FMS and learning gain. However, there are several cases, for example of learners with high FMS but low SMS. Or in an extreme case, a full FMS but zero SMS. This indicates that the learner can recompose a good concept map but may lack knowledge elaboration as reflected by their recomposing process. As SMS is the better predictor for learning gain than FMS, means that the process of knowledge elaboration plays an important aspect in the outcomes of learning gain. High correlation between SMS and FMS also depicts the relationship between knowledge elaboration process and map quality. This is one aspect that is potentially realized by SMS, in comparison with assessing the final concept map score realized by FMS. Therefore, it is a good approach for teachers to monitor learners and guide them to a better process in the concept map recomposition activity.

Thus far, there is no support for the recomposition process in KB concept mapping. However, if the sub-map recomposition behavior affects learning gain, then it is necessary to design a recomposition in-process support based on the sub-map structure. Consequently, some of the important future tasks include designing such support and evaluating its effectiveness. A more thorough investigation on the evaluation of knowledge elaboration and its relations with the concept mapping activity could also reveal more aspects of learner's understanding.

Conclusions and future works

Using the kit-build concept mapping with Sub-Map Scoring (SMS) demonstrates how the proposed method analyzes the learner's concept map recomposition process. This study's main contribution is its proposed technique for measuring the kit-building concept map recomposition process—the SMS, which examines the connected proposition sequences using the sub-maps defined by the teacher. The findings revealed that SMS, as well as FMS, has a high correlation with test scores and learning gain. This answered the first research question with a high relationship between SMS and learning gain.

When compared between pre-test score, SMS, and FMS in predicting the learning gain, SMS proved to be a better predictor. Therefore, it answered the second research question. Learners with high score of SMS would also have a high learning gain and vice versa. Kit-build concept mapping activity has also been proven to result in successful learning with no significant memory degradation in the student's comprehension of the learning material.

In addition, this study's findings revealed that the sub-map recomposition behavior impacts learning gain. Therefore, designing and providing in-process support for concept map recomposition based on the sub-map structure is imperative in the future works. Evaluating the process evaluation in the context of formative assessment could also reveal the benefits of the proposed method.

Abbreviations
KB: Kit-Build; FMS: Full Map Scoring; OOP: Object-Oriented Programming; SB: Scratch-Build; SMS: Sub-Map Scoring.

Acknowledgements
The authors would like to thank all the persons for their contributions, and Indonesia Endowment Funds for Education (LPDP) for supporting this study.

Authors' contributions
RR managed the experiments, developed the tool, analyzed the data, and wrote the manuscript. AP designed and developed the system and analyzed the data. BSA prepared and assisted the experiments. YH discussed and reviewed the analysis results. TH reviewed the experiment design, analysis results, and the manuscript. All authors read and approved the final manuscript.

Funding
This work was partially supported by JSPS KAKENHI Grant Number 19H04227 and Indonesia Endowment Funds for Education (LPDP).

Availability of data and materials
Availability of data and materials upon reasonable requests.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Informatics and Data Science, Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1 Kagamiyama, 739-8527, Higashihiroshima, Japan.

[2] Department of Information Technology, State Polytechnic of Malang, Jl. Soekarno Hatta No. 9, Lowokwaru, 65141, Malang, Indonesia.

[3] Department of Information System, Faculty of Computer Science, Universitas Brawijaya, Jl. Veteran No.8, Lowokwaru, 651145, Malang, Indonesia.

Received: 28 November 2022 Accepted: 18 August 2023
Published online: 1 January 2024 (Online First: 14 September 2023)

References

Alkhateeb, M., Hayashi, Y., Rajab, T., & 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. https://doi.org/10.12937/ejsise.14.13

Andoko, B. S., Hayashi, Y., Hirashima, T., & Asri, A. N. (2020). Improving English reading for EFL readers with reviewing kit-build concept map. *Research and Practice in Technology Enhanced Learning*, *15*, 7. https://doi.org/10.1186/s41039-020-00126-8

Bao, L. (2006). Theoretical comparisons of average normalized gain calculations. *American Journal of Physics*, *74*(10), 917–922. https://doi.org/10.1119/1.2213632

Bergman, E. M., de Bruin, A. B. H., Vorstenbosch, M. A. T. M., Kooloos, J. G. M., Puts, G. C. W. M., Leppink, J., Scherpbier, A. J. J. A., & van der Vleuten, C. P. M. (2015). Effects of learning content in context on knowledge acquisition and recall: A pretest-posttest control group design. *BMC Medical Education*, *15*, 133. https://doi.org/10.1186/s12909-015-0416-0

Cañas, A. J., Reiska, P., & Novak, J. D. (2016). Is my concept map large enough? *Communications in Computer and Information Science*, *635*, 128–143. https://doi.org/10.1007/978-3-319-45501-3_10

Chan, Z. C. Y. (2017). A qualitative study on using concept maps in problem-based learning. *Nurse Education in Practice*, *24*, 70–76. https://doi.org/10.1016/j.nepr.2017.04.008

Chiu, C.-H., & Lin, C.-L. (2012). Sequential pattern analysis: Method and application in exploring how students develop concept maps. *The Turkish Online Journal of Educational Technology*, *11*(1), 145–153.

Coletta, V. P., & Steinert, J. J. (2020). Why normalized gain should continue to be used in analyzing preinstruction and postinstruction scores on concept inventories. *Physical Review Physics Education Research*, *16*, 010108. https://doi.org/10.1103/PhysRevPhysEducRes.16.010108

Dwi Prasetya, D., Hirashima, T., & Hayashi, Y. (2020). Study on Extended Scratch-Build concept map to enhance students' understanding and promote quality of knowledge structure. *International Journal of Advanced Computer Science and Applications*, *11*(4), 144–153.

Elhelou, M.-W. A. (1997). The use of concept mapping in learning science subjects by Arab students. *Educational Research*, *39*(3), 311–317. https://doi.org/10.1080/0013188970390306

Ford, J. K., Smith, E. M., Weissbein, D. A., Gully, S. M., & Salas, E. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of Applied Psychology*, *83*(2), 218–233. https://doi.org/10.1037/0021-9010.83.2.218

Funaoi, H., Ishida, K., & Hirashima, T. (2011). Comparison of Kit-Build and Scratch-Build concept mapping methods on memory retention. In T. Hirashima et al. (Eds.), *Proceedings of the 19th International Conference on Computers in Education* (pp. 539–546). Asia-Pacific Society for Computers in Education.

Furtado, P. G. F., Hirashima, T., & Hayashi, Y. (2019). Reducing cognitive load during closed concept map construction and consequences on reading comprehension and retention. *IEEE Transactions on Learning Technologies*, *12*(3), 402–412. https://doi.org/10.1109/TLT.2018.2861744

Gerstner, S., & Bogner, F. X. (2009). Concept map structure, gender and teaching methods: An investigation of students' science learning. *Educational Research*, *51*(4), 425–438. https://doi.org/10.1080/00131880903354758

Gregoriades, A., Pampaka, M., & Michail, H. (2009). Assessing students' learning in MIS using concept mapping. *Journal of Information Systems Education*, *20*(4), 419–430.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74. https://doi.org/10.1119/1.18809

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*, 17. https://doi.org/10.1186/s41039-015-0018-9

Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, *77*(1), 95–111. https://doi.org/10.1002/sce.3730770107

Jonassen, D. H., Beissner, K., Yacci, M., & Beissner, K. (2013). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge*. Routledge. https://doi.org/10.4324/9780203052563

Kalyuga, S. (2009). Knowledge elaboration: A cognitive load perspective. *Learning and Instruction*, *19*(5), 402–410. https://doi.org/10.1016/j.learninstruc.2009.02.003

Kinchin, I. M., & Alias, M. (2005). Exploiting variations in concept map morphology as a lesson-planning tool for trainee teachers in higher education. *Journal of In-Service Education*, *31*(3), 569–592.

Kinchin, I. M., Hay, D. B., & Adams, A. (2000). How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. *Educational Research*, *42*(1), 43–57. https://doi.org/10.1080/001318800363908

Lee, M., & Fortune, A. E. (2013). Patterns of field learning activities and their relation to learning outcome. *Journal of Social Work Education*, *49*(3), 420–438. https://doi.org/10.1080/10437797.2013.796786

Lee, W., Chiang, C.-H., Liao, I.-C., Lee, M., Chen, S.-L., & Liang, T. (2013). The longitudinal effect of concept map teaching on critical thinking of nursing students. *Nurse Education Today*, *33*(10), 1219–1223.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process*. *British Journal of Educational Psychology*, *46*(1), 4–11. https://doi.org/https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

Marx, J. D., & Cummings, K. (2007). Normalized change. *American Journal of Physics*, *75*(1), 87–91. https://doi.org/10.1119/1.2372468

McClure, J. R., & Bell, P. E. (1990). *Effects of an environmental education-related STS approach instruction on cognitive structures of preservice science teachers*. Pennsylvania State University.

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. https://doi.org/10.1002/(SICI)1098-2736(199904)36:4<475::AID-TEA5>3.0.CO;2-O

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358–368. https://doi.org/10.1037/0022-0663.91.2.358

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, *27*(10), 937–949. https://doi.org/10.1002/tea.3660271003

Novak, J. D., & Symington, D. (1982). Concept mapping for curriculum development. *VIER Bulletin*, *48*, 3–11.

Novak, J. D., Gowin, D. B., & Kahle, J. B. (1984). *Learning how to learn*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173469

Pailai, J., Wunnasri, W., Hayashi, Y., & Hirashima, T. (2018). Correctness-and confidence-based adaptive feedback of Kit-Build concept map with confidence tagging. In C. Penstein Rosé et al. (Eds.), *Artificial Intelligence in Education. AIED 2018. Lecture Notes in Computer Science, vol 10947* (pp. 395–408). Springer, Cham. https://doi.org/10.1007/978-3-319-93843-1_29

Pailai, J., Wunnasri, W., Yoshida, K., Hayashi, Y., & Hirashima, T. (2017). The practical use of Kit-Build concept map on formative assessment. *Research and Practice in Technology Enhanced Learning*, *12*, 20. https://doi.org/10.1186/s41039-017-0060-x

Papajohn, D. (2002). Concept mapping for rater training. *TESOL Quarterly*, *36*(2), 219–233. http://www.jstor.org/stable/3588333

Plotnick, E. (1997). *Concept mapping: A graphical system for understanding the relationship between concepts*. ERIC Clearinghouse on Information and Technology.

Roberts, L. (1999). Using concept maps to measure statistical understanding. *International Journal of Mathematical Education in Science and Technology*, *30*(5), 707–717. https://doi.org/10.1080/002073999287707

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, *33*(6), 569–600. https://doi.org/https://doi.org/10.1002/(SICI)1098-2736(199608)33:6<569::AID-TEA1>3.0.CO;2-M

Sadita, L., Hirashima, T., Hayashi, Y., Wunnasri, W., Pailai, J., Junus, K., & Santoso, H. B. (2020). Collaborative concept mapping with reciprocal kit-build: A practical use in linear algebra course. *Research and Practice in Technology Enhanced Learning*, *15*, 19. https://doi.org/10.1186/s41039-020-00136-6

Schneider, S., Krieglstein, F., Beege, M., & Rey, G. D. (2021). How organization highlighting through signaling, spatial contiguity and segmenting can influence learning with concept maps. *Computers and Education Open*, *2*, 100040. https://doi.org/10.1016/j.caeo.2021.100040

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review*, *30*(3), 679–701. https://doi.org/10.1007/s10648-018-9435-9

Srivastava, A., Srivastava, N., & Chandrasekharan, S. (2021). Quality of concept maps is affected by map building strategies. *Biochemistry and Molecular Biology Education*, *49*(1), 129–139. https://doi.org/https://doi.org/10.1002/bmb.21412

Sundararajan, N., Adesope, O., & Cavagnetto, A. (2018). The process of collaborative concept mapping in kindergarten and the effect on critical thinking skills. *Journal of STEM Education*, *19*(1).

Van Rossum, E. J., & Schenk, S. M. (1984). The relationship between learning conception, study strategy and learning outcome. *British Journal of Educational Psychology*, *54*(1), 73–83. https://doi.org/10.1111/j.2044-8279.1984.tb00846.x

Vanides, J., Yin, Y., Tomita, M., & Ruiz-Primo, M. A. (2005). Using concept maps in the science classroom. *Science Scope*, *28*(8), 27–31.

Weinstein, C., & Mayer, R. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315–327). Macmillan.

West, D. C., Pomeroy, J. R., Park, J. K., Gerstenberger, E. A., & Sandoval, J. (2000). Critical thinking in graduate medical education: A role for concept mapping assessment? *JAMA*, *284*(9), 1105–1110. https://doi.org/10.1001/jama.284.9.1105

Wunnasri, W., Pailai, J., Hayashi, Y., & Hirashima, T. (2018a). Reciprocal Kit-Build concept map: An approach for encouraging pair discussion to share each other's understanding. *IEICE Transactions on Information and Systems*, *E101D*(9), 2356–2367. https://doi.org/10.1587/transinf.2017EDP7420

Wunnasri, W., Pailai, J., Hayashi, Y., & Hirashima, T. (2018b). Validity of Kit-Build method for assessment of learner-build map by comparing with manual methods. *IEICE Transactions on Information and Systems*, *E101D*(4), 1141–1150. https://doi.org/10.1587/transinf.2017EDP7177

Yamasaki, K., Fukuda, H., Hirashima, T., & Funaoi, H. (2010). Kit-Build concept map and its preliminary evaluation. In S. L. Wong et al. (Eds.), *Proceedings of the 18th International Conference on Computers in Education* (pp. 290–294). Asia-Pacific Society for Computers in Education.

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. In K. Yoshida et al. (Eds.), *Proceedings of the 21st International Conference on Computers in Education* (pp. 892–901). Asia-Pacific Society for Computers in Education.

Zheng, L., Li, X., Zhang, X., & Sun, W. (2019). The effects of group metacognitive scaffolding on group metacognitive behaviors, group performance, and cognitive load in computer-supported collaborative learning. *Internet and Higher Education*, *42*, 13–24. https://doi.org/10.1016/j.iheduc.2019.03.002

Zheng, L., Niu, J., & Zhong, L. (2022). Effects of a learning analytics-based real-time feedback approach on knowledge elaboration, knowledge convergence, interactive relationships and group performance in CSCL. *British Journal of Educational Technology*, *53*(1), 130–149. https://doi.org/10.1111/bjet.13156

Publisher's Note

The Asia-Pacific Society for Computers in Education (APSCE) remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.