Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Trial-and-Error Modifications as Cognitive Compensation: Analyzing the Concept Mapping Process in Reconstruction Tasks

Kodai WATANABE[a*], Yuma CHOKI[a], Ridwan RISMANTO[b], Shimpei MATSUMOTO[c] & Tsukasa HIRASHIMA[a]

[a]*Graduate School of Advanced Science and Engineering, Hiroshima University, Japan*
[b]*Department of Information Technology, State Polytechnic of Malang, Indonesia*
[c]*Faculty of Information and Communications, Hiroshima Institute of Technology, Japan*
*kodaiwatanabe10@gmail.com

Abstract: The importance of trial-and-error modifications during concept mapping has been widely acknowledged, especially as an opportunity for learners to engage in metacognitive and higher-order thinking. However, it has been difficult to systematically analyze such modification activities due to the freedom learners have in creating and labeling nodes and links. This study investigates modification processes using the framework of the reconstruction-based concept map, in which a teacher-constructed map is deconstructed into its component nodes and links first, and then learners reconstruct the map using only these provided components. This setting enables the diagnosis of misconceptions—based on deviations from the teacher-constructed map—and consistent, comparable analysis of the mapping process across learners. From the learner's perspective, prior studies have shown its effectiveness in promoting meaningful learning and higher-order thinking. We analyzed three datasets collected from university students enrolled in an object-oriented programming course. Each dataset corresponds to a different topic, with learners individually reconstructing maps using a Web-based application. The analysis focused on two process indicators: the number of false propositions (propositions not matching the teacher-constructed map) and the number of proposition changes, defined as modifying an existing proposition by reconnecting its link to a node. Regression analyses showed that while proposition changes alone had little correlation with map scores (an indicator of map quality), combining proposition changes with false propositions yielded a strong and significant correlation. Further comparison between the high-scoring group and the low-scoring group revealed no significant differences in the frequency or precision of proposition changes. However, high-scoring learners showed significantly higher rates of successful corrections (changes from a false proposition to a correct proposition) and a better recall rate of false propositions. These findings suggest that a form of cognitive compensation may play a key role in concept mapping performance—namely, the ability to correct earlier errors despite incomplete understanding. For example, high-scoring learners still created false propositions, but were more effective at correcting them. The study emphasizes the importance of supporting proposition modification activities, particularly for low-scoring learners, and provides insights into the design of scaffolding mechanisms in digital concept mapping environments.

Keywords: Concept map, reconstruction-based concept map, process analysis, trial-and-error modifications, cognitive compensation

1. Introduction

In studies on human learning, it has been emphasized that analyzing not only the final products but also the processes that generate them is essential (Ford et al., 1998; Lee & Fortune, 2013). Concept mapping is one such learning activity whose effectiveness has been widely recognized, and numerous studies have evaluated concept maps as products (Novak & Gowin, 1984; McClure & Bell, 1990; McClure et al., 1999). A concept map is a diagrammatic representation that depicts the semantic structure by linking propositions, each consisting of two concepts (nodes) connected by a labeled link. Propositions share nodes, and thereby form an integrated semantic structure. The importance of the concept mapping process itself has also been emphasized. For instance, Jablokow et al. (2015) noted that it can reveal detailed aspects of learners' understanding. Cañas et al. (2017) further argued that reviewing and modifying a map during its construction is a key activity that fosters metacognition and higher-order thinking.

However, process analyses to date have been limited to examining aspects such as the order in which nodes and links are placed (Srivastava et al., 2021) and the number of modification activities (Ching & Hsu, 2011). There have also been attempts to diagnose concept maps using natural language processing techniques automatically, yet none of these methods can explicitly determine the validity of individual propositions (Bhatia et al., 2021; Bleckmann & Friege, 2023). Consequently, no prior studies—whether qualitative or quantitative—have empirically analyzed the content of individual modifications to the propositions created. This limitation stems from the fact that, in standard concept mapping, learners are free to add or remove nodes and links, and the terminology and phrasing they use can vary from one learner to another. In such an environment, even if each learner's modification history could be tracked, it would be challenging to consistently compare the results across learners.

From a theoretical perspective, the process of modifying a concept map can be seen as an attempt to align a learner's internal representation—mental models or conceptual structures—with an external representation provided in the learning environment. External representations convey both elemental meaning (meaning inherent in individual components) and structural meaning (meaning arising from their organization). The additional meaning generated through the integration of these components, termed constructed meaning, emerges only when learners actively reorganize the structure (Hirashima & Watanabe, 2025). Proposition modifications in concept mapping, therefore, can be interpreted as opportunities for refining constructed meaning through the manipulation of external representations, potentially triggering metacognitive reflection and conceptual change. These theoretical insights highlight the significance of proposition modifications, yet their empirical analysis remains constrained by the limitations described above.

The Reconstruction-based Concept Map (RCM) proposed by the authors in previous work (Hirashima et al., 2015; Hirashima, 2024) may offer a solution to the above problem. In an RCM task, the teacher first constructs a concept map—referred to in this paper as the shared understanding map—that represents the understanding to be shared with all learners, such as prerequisite knowledge for the next lesson. This map is then deconstructed into its constituent nodes and links, which are provided to learners as components. Learners reconstruct the shared understanding map by connecting these given components to form the learner map. In this paper, the operation of connecting components is termed the map construction activity (concept mapping). A modification activity is defined as (1) disconnecting a link in an existing proposition, and (2) reconnecting it to a node to form a proposition. A modification is recognized when both actions occur, although they do not need to be performed consecutively. From the perspective of a learning activity, prior studies have shown that reconstruction of concept maps can promote meaningful learning (Pailai et al., 2017) and higher-order thinking (Nurmaya et al., 2023). These findings suggest that RCM activities are worth analyzing from a process perspective, as they provide learners with adequate opportunity for trial-and-error, even within the reconstruction of predetermined components.

Rismanto et al. (2024) conducted a process analysis utilizing the aforementioned characteristics of RCMs. They divided a concept map into several semantically grouped submaps and examined whether the sequential creation of propositions belonging to the same submap—i.e., viewing a concept map in terms of semantic groupings—affected comprehension of the learning target. However, their analysis focused on the order in which propositions consistent with the shared understanding map were created. It did not examine how the map score—the degree of agreement between the learner map and the shared understanding map—was affected by modification activities during the process.

In light of the above, this paper poses the research question: *How do learners' modification activities affect map scores in RCMs?* It reports on an analysis of datasets obtained from three lessons in which RCMs were used.

2. Concept Map and Its Reconstruction

2.1 Concept Map

Concept maps, which visually represent the relationships between concepts, serve two main purposes: as a tool for evaluating learners' understanding and as a learning tool for deepening that understanding (Cañas et al., 2023). Regarding their use as an evaluation tool, numerous methods have been proposed that assess the quality of the constructed map. Novak and Gowin (1984) proposed a method focusing on structural features of the map, such as the hierarchical arrangement of concepts and the links connecting them. In contrast, McClure and Bell (1990) proposed a method that emphasizes the semantic accuracy of each proposition. McClure et al. (1999) conducted a comparative analysis of six evaluation methods, including the above, from the perspectives of reliability and validity, and systematically organized these methods.

From the perspective of its use as a learning tool, concept mapping has been recognized for its effectiveness in promoting learners' metacognition and higher-order thinking. Regarding higher-order thinking, Anderson et al. (2001) clarified its constituent elements based on Bloom's taxonomy, a classification of educational objectives. Specifically, remember, understand, and apply are categorized as lower-order thinking. In contrast, analyze, evaluate, and create are regarded as higher-order thinking, which involves deeper cognitive processing. As a rationale for the promotion of higher-order thinking through concept mapping, Cañas et al. (2017) noted that during proposition construction—such as adding or removing nodes and links or changing labels—learners continually engage in meta-evaluation, assessing whether the operations they intend to perform appropriately represent their understanding of the learning topic.

The RCMs used in this study function as learning activities in classroom settings and thus possess the characteristics of learning tools. At the same time, because this study analyzes the relationship between indicators of modification activities during concept mapping and the quality of the final maps (map scores), it also addresses their use as evaluation tools. Note that, although modification activities have been suggested to be associated with higher-order thinking, the present study prioritizes quantitatively capturing their basic characteristics. Verification of their relationship with higher-order thinking is therefore left for future work.

2.2 Reconstruction-Based Concept Map

In standard concept mapping, learners freely create nodes and links based on their understanding and combine them to form propositions. Consequently, even for the same learning topic, different learners may use different nodes and links. For example, in a concept map about Japan, some learners might represent the concept of Japan as "Japan," while others might use "Our country." Moreover, a learner who knows about Japan's main imports but not its exports may omit the corresponding propositions. Therefore, evaluating the constructed concept maps and their mapping processes requires addressing the diversity of learners' expressions and content. Since this cannot be done without human judgment, it places a considerable burden on teachers.

In contrast, in RCMs, as illustrated in Figure 1, (1) the teacher constructs a shared understanding map and (2) deconstructs it into nodes and links. These components are subsequently provided to the learner. Then, (3) the learner reconstructs the concept map using only the provided components—a defining feature of RCMs. Because the shared understanding map and the learner map share the same nodes and links, (4) the two maps can be superimposed, and each link examined to determine whether the proposition formed by its connected nodes is identical to that in the shared understanding map. Agreement or mismatch for each proposition is automatically detected, forming the basis for calculating the map score. This comparison can also be visualized, as in the lower left of Figure 1. The map score is defined as the ratio of correct propositions—those present in both maps—to the total propositions in the shared understanding map. In the example, the shared understanding map contains three propositions, and the only correct proposition in the learner map is “Japan -> Capital city -> Tokyo,” yielding a score of 1 / 3 ≈ 0.33. In the visualization, the false proposition—i.e., the one present only in the learner map—“Japan -> Main imports -> Automobiles” is shown with a solid line, while the missing propositions—those present only in the shared understanding map—“Japan -> Main imports -> Crude oil” and “Japan -> Main exports -> Automobiles” are shown with dashed lines. This shared-component structure of RCMs provides a robust basis for quantitatively analyzing learners’ concept mapping processes.

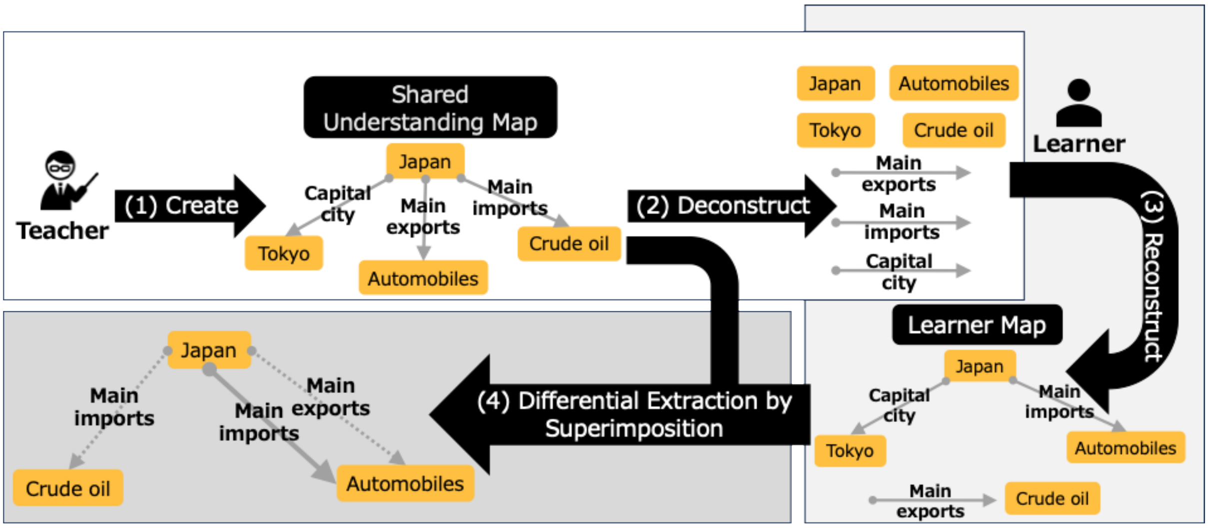

Figure 1. Framework for Reconstruction-based Concept Map.

3. Classroom Practices and Data Sets

In this study, we analyzed datasets collected by one of the co-authors during an object-oriented programming course for second-year undergraduate students at a university in Indonesia, as part of a teaching practice designed to examine the learning effects of RCMs. The practice focused on key topics in object-oriented programming: the first lesson covered encapsulation, the second covered inheritance, and the third covered polymorphism. All three lessons were conducted in the same format. Hereafter, each topic is abbreviated as En, In, and Po, respectively.

The analysis targeted 30 learners who participated in all three lessons. All had completed a basic programming course in their first year, but this was their first exposure to object-oriented programming. None had any prior knowledge or experience with concept mapping. In each lesson, the teacher began with a 15-minute lecture on the topic. Then, to deepen their understanding, the learners used a Web-based application to reconstruct a concept map for 15 minutes. During the concept mapping, they were not allowed to refer to the teaching materials. Although the Web-based application includes a feedback function based on automatic diagnosis for learner maps described in the previous chapter, it was used without feedback in this practice. This was intended to replicate the conditions of standard concept mapping.

The shared understanding maps used in these practices were constructed by one of the co-authors, then reviewed and revised based on feedback from other members of the teaching team. Table 1 presents the counts of nodes and links, as well as the summary statistics of map scores for each shared understanding map. Although each dataset covered the same learners (n = 30), learners who were deemed not to have engaged in appropriate learning activities were excluded from the analysis. Specifically, exclusion was applied when either the number of propositions or the time to reconstruct the map was identified as an outlier. The number of propositions refers to the cumulative count of propositions created by a learner—regardless of correctness or repetition—which serves as an indicator of learner activity in RCM tasks. Outliers were defined as values greater than the third quartile + 1.5 × the interquartile range, or smaller than the first quartile − 1.5 × the interquartile range. As a result, the number of learners analyzed was n = 28 for En, n = 29 for In, and n = 26 for Po.

Table 1. *Shared Understanding Maps: Node/Link Counts and Map Score Summary Statistics*

Map	Nodes	Links	Median	Mean	*SD*
En	14	15	.63	.62	.34
In	12	13	.92	.77	.28
Po	11	10	.90	.85	.17

4. Quantitative Analysis of the Concept Mapping Process

4.1 The Impact of Proposition Changes and False Propositions on Map Scores

In this study, we first examined the number of proposition changes as an indicator of modification activity in RCMs. This indicator is defined as the cumulative count of modification activities made by a learner. For example, if a learner created the proposition "Japan -> Main imports -> Crude oil" and then changed the link destination to "Automobiles," this would be counted as one proposition change. Table 2 presents the mean and standard deviation of the proposition changes, along with the results of multiple regression analysis described below. Due to page limitations, the results of the single regression analysis—conducted with the map score as the dependent variable—are omitted; however, no strong correlations were found in any of the maps. In this paper, we adopt the widely accepted criterion that $r \geq .70$ indicates a strong correlation, and correspondingly interpret an R^2 (or *adjusted* R^2 in the case of multiple regression) ≥ .50 in regression analyses reflects a strong correlation between the explanatory variables and the dependent variable.

Table 2. *Mean and Standard Deviation of the Number of Proposition Changes and False Propositions / Results of Multiple Regression Analysis on Map Scores*

Map	Proposition Changes	False Propositions	Linear Regression	*VIF*
En	3.50 (*SD* = 3.46)	6.25 (*SD* = 4.44)	.59 ($p < .001$)	1.42
In	4.03 (3.20)	5.76 (4.35)	.87 ($p < .001$)	1.40
Po	2.19 (2.53)	3.27 (2.76)	.88 ($p < .001$)	2.83

Since proposition changes occur when a learner judges that there are false propositions in the map, the map score cannot be explained solely by the proposition changes without also considering the number of false propositions. For example, even if the number of proposition changes is low, a small number of false propositions in the map may indicate that the modification activity was sufficient. Conversely, if there are many false propositions, the same number of modifications could be judged as an inadequate response. Therefore, since the significance of a modification activity is considered to vary depending on the number of false propositions, this study defined the latter as a separate indicator and conducted a combined analysis using both. The number of false propositions is defined as the cumulative count of false propositions created by a learner, regardless of repetition. For example, in the earlier case of changing from “Japan -> Main imports -> Crude oil” to “Japan -> Main imports -> Automobiles,” the proposition after the change is a false proposition. In such cases, the false proposition is counted in addition to the proposition change. Table 2 presents the mean and standard deviation of the false propositions.

Table 2 shows the results of multiple regression analysis using proposition changes and false propositions as explanatory variables, with the map score as the dependent variable. The table reports the *adjusted* R^2 and the p-value from the overall F-test for the regression model. In this paper, we set the significance level at 5%. The *Variance Inflation Factor (VIF)* values are shown in the "*VIF*" column; in all cases, the *VIF* was well below the common multicollinearity threshold of 10, indicating no multicollinearity among the explanatory variables. In this analysis, strong correlations (*adjusted* $R^2 \geq .50$) were observed for all three maps. Although detailed results, such as regression equations and single regression analysis, are omitted due to page limitations, the sign of the regression coefficients indicates that false propositions negatively affect the map score, whereas proposition changes have a positive effect. In single regression analysis, false propositions, when considered alone, showed strong correlations with the map score for In and Po maps. This may be attributed to a ceiling effect: since the sum of the mean and standard deviation of the map scores shown in Table 1 exceeded 1 (maximum map score), the weight of each false proposition on the map score was relatively large. Intuitively, the more false propositions learners created, the less likely they were to correct all of them by the end, which may have further reinforced this effect. While the present analysis does not directly examine this possibility, related aspects of these processes are analyzed in sections 4.2 and 4.3, including how learners successfully corrected false propositions and the proportion of false propositions that remained unaddressed.

These results quantitatively address the research question of this study—*How do learners' modification activities affect map scores in RCMs?*—by demonstrating that modification activities do influence map scores.

4.2 Analysis of Proposition Changes

The analysis in the previous section provides an answer to the research question. Nevertheless, this finding alone does not explain how these two indicators capture the nature of learners' modification activities or how each is related to the map score.

Of the two indicators, false propositions are inherently negatively related to the map score. As shown in Table 3, when learners were divided into two groups based on whether their map score was at or above the median, the high-scoring group (High Group) had significantly fewer false propositions than the low-scoring group (Low Group). Since some variables did not meet the assumption of normality, all comparisons of means in this study were conducted using the Mann-Whitney U test. The effect sizes (r) are also reported in Table 3 as a supplement to the p-values. The median was chosen as the cutoff because a ceiling effect was observed for the In and Po maps.

Table 3. *Difference in the Mean Number of False Propositions between the High-Scoring Group (High Group) and the Low-Scoring Group (Low Group)*

Map	Group	False Propositions	p and r
En	High Group (n = 14)	4.21 (SD = 3.60)	p = .02,
	Low Group (n = 14)	8.29 (4.36)	r = .46
In	High Group (n = 15)	3.20 (2.78)	p < .01,
	Low Group (n = 14)	8.50 (4.09)	r = .61
Po	High Group (n = 16)	2.38 (2.50)	p = .03,
	Low Group (n = 10)	4.70 (2.67)	r = .44

In contrast, proposition changes differ in their effects on the map score: changes targeting correct propositions may lead to a decrease in the map score, whereas those targeting false propositions may lead to an increase in the map score. In this section, we report the results of our investigation into the types of proposition changes made by the high and low groups.

First, we examined whether there were significant differences in the mean number of proposition changes between groups, as shown in Table 4. None of the results were significant, suggesting that differences in map scores were influenced more by the qualitative aspects of proposition changes than by their quantitative aspects (i.e., the number of proposition changes). In this study, the qualitative aspects of proposition changes are defined in terms of two factors: (1) the appropriateness of the change target—whether it was applied to a correct or a false proposition—and (2) the success or failure of the change—whether a change targeting a false proposition resulted in its correction to a correct proposition. Based on these two factors, all proposition changes can be classified into combinations of correctness, such as "F-C" (false to correct) or "F-F" (false to false), allowing for a comprehensive understanding of proposition changes.

Table 4. *Difference in the Mean Number of Proposition Changes between the High Group and the Low Group*

Map	Group	Proposition Changes	*p* and *r*
En	High	3.93 (*SD* = 4.29)	$p = .96$, $r = .01$
	Low	3.07 (2.46)	
In	High	3.87 (3.44)	$p = .61$, $r = .10$
	Low	4.21 (3.04)	
Po	High	2.19 (2.79)	$p = .80$, $r = .05$
	Low	2.20 (2.20)	

To analyze (1) the appropriateness of the change target, Table 5 presents a breakdown of changes into those targeting correct propositions and those targeting false propositions. The table reports the mean value for each group, with percentages shown in parentheses, indicating each type's share of the total proposition changes within the group. The results indicate that changes targeting false propositions consistently accounted for more than 70% of proposition changes in both groups across all three maps. In other words, regardless of whether the map score is high or low, it indicates a generally high level of precision in proposition changes.

Table 5. *Targets of Proposition Changes in the High Group and the Low Group*

Map	Group	Changes Targeting Correct Propositions	Changes Targeting False Propositions
En	High	1.07 (27.27%)	2.86 (72.73%)
	Low	0.36 (11.63%)	2.71 (88.37%)
In	High	0.73 (18.97%)	3.13 (81.03%)
	Low	1.14 (27.12%)	3.07 (72.88%)
Po	High	0.31 (14.29%)	1.88 (85.71%)
	Low	0.60 (27.27%)	1.60 (72.73%)

To analyze (2) the success or failure of the change, we categorized changes targeting false propositions into F-C and F-F. Table 6 presents the total number of each type of proposition change for learners in each group. The table also reports the *p*-values and effect sizes (*w*) from chi-square tests comparing the distribution of proposition changes between groups within each map, along with the results of residual analyses. In the table, "▲" indicates that the value for the corresponding group is significantly higher than the expected frequency, while "▽" indicates that it is significantly lower. The results show a clear contrast: the high group made significantly more F-C changes in the En and In maps, while the low group made significantly more F-F changes. This pattern may reflect a cognitive compensation process (diSessa, 1993), in which learners improve their final performance by drawing on partial but accurate knowledge to correct earlier errors. In this study, high-scoring learners sometimes created false propositions but successfully corrected them, suggesting that they were effectively leveraging cognitive compensation processes during concept mapping.

Table 6. *Distribution of the Success or Failure of Proposition Changes Targeting False Propositions between the High Group and the Low Group, with Results of Chi-Square Tests and Residual Analyses*

Map	Group	F-C (false to correct)	F-F (false to false)	*p* and *w*
En	High	28▲	12▽	$p < .001$, $w = .38$
	Low	11▽	27▲	
In	High	33▲	14▽	$p < .001$, $w = .47$
	Low	9▽	34▲	
Po	High	21	9	$p = .31$, $w = .15$
	Low	8	8	

This difference can also be interpreted within the framework of Reconstruction-Based Learning (Hirashima & Watanabe, 2025), which views learning as the refinement of constructed meaning—the additional meaning generated from the organization of elemental and structural meaning—through the manipulation of external representations. In this framework, the high group's greater success in F-C changes and higher false proposition coverage rate (a metric described in the following section) indicate more effective engagement in Stage 2 (difference detection) and Stage 3 (conceptual clarification and completion), leading to better alignment between internal and external representations. In contrast, the low group's lower F-C rates suggest difficulty in completing this refinement process, even when engaging in a comparable number of modification activities.

A similar tendency was also observed in the Po map, where the majority of proposition changes made by the high group were F-C changes (exact binomial test: $p = .04$, $w = .40$), while the low group did not show this pattern. Taken together, these findings suggest that although both groups make a similar number of proposition changes, and most of these target false propositions, it is the successful correction of those false propositions that differentiates high-scoring learners from low-scoring ones.

4.3 Analysis of False Proposition Coverage Rate

In the previous section, we examined proposition changes made by learners and found a significant difference between the high and low groups in terms of the success or failure of changes targeting false propositions (Table 6). However, those analyses did not assess the extent to which false propositions remained unchanged. For example, if learner A created a total of 10 false propositions during the concept mapping and made 9 proposition changes, nearly all of their false propositions would have been changed. By contrast, if learner B also made 9 proposition changes but created 20 false propositions in total, a large number would have remained unchanged.

In this section, we address the above issue by introducing a metric termed the false proposition coverage rate. This metric is defined as the ratio of the number of changes targeting false propositions to the number of false propositions, and is conceptually equivalent to the recall rate of proposition changes.

Table 7 presents the results of examining whether there were significant differences in false proposition coverage rates between groups. Learners with no false propositions were excluded from the analysis, as the coverage rate could not be calculated due to division by zero.

The test results showed significant differences between the high and low groups across all three maps. Although the detailed data are omitted here due to page limitations, the low group spent as much or more time reconstructing the map than the high group, indicating that time use was not an issue. In other words, although the low group spent enough time, about 70% of their false propositions tended to remain unchanged.

Table 7. *Difference in the Mean of False Proposition Coverage Rates between the High Group and the Low Group*

Map	Group	False Proposition Coverage Rate	p and r
En	High (n = 13)	.69 (SD = .34)	p < .01,
	Low (n = 14)	.32 (.21)	r = .54
In	High (n = 13)	.99 (.03)	p < .001,
	Low (n = 13)	.36 (.20)	r = .91
Po	High (n = 11)	.68 (.37)	p < .05,
	Low (n = 10)	.28 (.26)	r = .58

5. Conclusion

In this paper, we addressed the lack of empirical studies—both qualitative and quantitative—that examine individual concept map modification activities, despite their recognized importance in promoting learning. Drawing on data from classroom practices involving RCMs—which are well-suited for quantitative analysis—we examined how indicators of learners' modification activities relate to their map scores.

As a result, we found that (1) the number of proposition changes alone was not strongly correlated with the map score; however, multiple regression analysis that included both the number of proposition changes and false propositions yielded a strong correlation across all three datasets. Furthermore, when learners were divided into two groups based on whether their map scores were at or above the median, we found that (2) the number of proposition changes did not differ significantly between the two groups, and (3) the majority of these changes targeted false propositions, indicating a generally high level of precision for proposition changes. However, (4) while high-scoring learners were able to successfully correct false propositions—an ability interpretable as cognitive compensation—low-scoring learners often failed to do so. In addition, the analysis of the false proposition coverage rate revealed that (5) many false propositions remained unchanged by the end of the concept mapping, suggesting a tendency toward a low recall rate of proposition changes in the low group.

Given that engaging in modification provides learners with valuable learning opportunities, the creation of a false proposition does not necessarily require immediate correction. However, the low group's failure to correct false propositions, together with their low false proposition coverage rate, are issues that cannot be overlooked. These findings highlight the importance of providing targeted support for proposition changes in this group. In future work, we plan to analyze common patterns underlying failures to correct false propositions and identify factors contributing to low coverage rates, with the aim of developing effective support strategies.

This study also has several limitations. First, the relatively small scale of the present analysis may limit statistical power; therefore, conducting studies on larger and more diverse datasets will be an important next step. Second, because RCMs involve constructing concept maps solely from predetermined components, the results cannot be directly generalized to standard, free-form concept mapping. Nevertheless, our findings underscore the importance of considering both modification activities and false proposition coverage when analyzing more open-ended mapping processes. Building on this insight, future work will extend the analysis to the standard concept mapping tasks and examine whether the observed patterns hold in those contexts.

Acknowledgements

This work was supported by JST SPRING, Grant Number JPMJSP2132.

References

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (Eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Addison Wesley Longman.

Bhatia, S., Bhatia, S., & Ahmed, I. (2021). Automated waterloo rubric for concept map grading. *IEEE Access*, *9*(1), 148590-148598. https://doi.org/10.1109/ACCESS.2021.3124672

Bleckmann, T., & Friege, G. (2023). Concept Maps for Formative Assessment: Creation and Implementation of an Automatic and Intelligent Evaluation Method. *Knowledge Management & E-Learning*, *15*(3), 433-447. https://doi.org/10.34105/j.kmel.2023.15.025

Cañas, A. J., Reiska, P., & Möllits, A. (2017). Developing higher-order thinking skills with concept mapping: A case of pedagogic frailty. *Knowledge management & e-learning*, *9*(3), 348-365. https://doi.org/10.34105/j.kmel.2017.09.021

Cañas, A. J., Reiska, P., & Shvaikovsky, O. (2023). Improving Learning and Understanding through Concept Mapping. *Knowledge Management & E-Learning*, *15*(3), 369-380. https://doi.org/10.34105/j.kmel.2023.15.021

Ching, Y.-H., & Hsu, Y.-C. (2011). Design-grounded assessment: A framework and a case study of Web 2.0 practices in higher education. *Australasian Journal of Educational Technology*, *27*(5), 781-797. https://doi.org/10.14742/ajet.931

diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, *10*(2-3), 105-225. https://doi.org/10.1080/07370008.1985.9649008

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

Hirashima, T. (2024). Formative assessment and meaningful learning with concept mapping through recomposition. *Information and Technology in Education and Learning*, *4*(1), Inv-p003. https://doi.org/10.12937/itel.4.1.Inv.p003

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), Article No. 17. https://doi.org/10.1186/s41039-015-0018-9

Hirashima, T., & Watanabe, K. (2025). Recomposition based learning for promoting structural understanding - From reconstruction of external representations to recomposition of internal representation -. In Jiang, B. et al. (Eds.), *Proceedings of the 33rd International Conference on Computers in Education* (in press). Asia-Pacific Society for Computers in Education.

Jablokow, K. W., DeFranco, J. F., Richmond, S. S., Piovoso, M. J., & Bilén, S. G. (2015). Cognitive style and concept mapping performance. *Journal of Engineering Education*, *104*(3), 303-325. https://doi.org/10.1002/jee.20076

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

McClure, J. R., & Bell, P. E. (1990). *Effects of an Environmental Education-Related STS Approach Instruction on Cognitive Structures of Preservice Science Teachers*. The 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: The Official Journal of the National Association for 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

Novak, J. D., & Gowin, D. B. (1984). *Learning How to Learn*. Cambridge University Press.

Nurmaya, Pinandito, A., Hayashi, Y., & Hirashima, T. (2023). Promoting students' higher order thinking with concept map recomposition. *IEICE Transactions on Information and Systems*, *E106.D*(8), 1262-1274. https://doi.org/10.1587/transinf.2022EDP7203

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*(1), Article No. 20. https://doi.org/10.1186/s41039-017-0060-x

Rismanto, R., Pinandito, A., Andoko, B. S., Hayashi, Y., & Hirashima, T. (2024). Evaluating the kit-build concept mapping process using sub-map scoring. *Research and Practice in Technology Enhanced Learning*, *19*(1), Article No. 21. https://doi.org/10.58459/rptel.2024.19021

Srivastava, A., Srivastava, N., & Chandrasekharan, S. (2020). Quality of concept maps is affected by map building strategies. *Biochemistry and Molecular Biology Education*, *49*(1), 129-139. https://doi.org/10.1002/bmb.21412