EMBEDDING CONCEPT MAPPING INTO UNIVERSITY MATHEMATICS: COMPARISON AND VALIDATION OF MARKING RUBRICS

Inae Jeong[a], Tanya Evans[a]

Presenting Author: Inae Jeong (ijeo175@aucklanduni.ac.nz)
[a]Department of Mathematics, University of Auckland, Auckland CBD, Auckland, New Zealand

KEYWORDS: Concept Map, Concept Map Evaluation, Assessment, Mathematics Education, Undergraduate Mathematics

ABSTRACT

Concept mapping is a visual way of presenting a group of related abstract concepts and identifying relationships between them by connecting related concepts with directed arrows that specify relationships. In the last few decades, concept mapping has become a popular research and educational tool. However, despite its extensive usage, not much research has been done in designing methods to evaluate concept mapping tasks and their validation. Moreover, very little has been reported about concept mapping usage in mathematics education. In this study, university students (N=260) in a large undergraduate mathematics course (for non-mathematics majors) were assigned to construct a concept map for *Vector Space*, which they had studied in the course. This research investigated various ways to evaluate students' concept mapping activity by comparing four rubrics. Using multiple linear regression to predict final exam outcomes, we were able to identify the best rubric for assessing student concept mapping. We found that the most important aspect in assessing concept mapping tasks is the inverse ratio between the number of concepts and the number of relationships between them presented in student work. This finding informs practical recommendations for implementing concept mapping activity in mathematics courses that we present at the end of the paper, together with a call for future research to investigate the causal relationship between the use of concept mapping and learning outcomes. It is of great interest to find out whether increasing the amount of concept mapping activity in a mathematics course would enhance student conceptual understanding.

INTRODUCTION

Origin and description of concept maps

Concept maps were first introduced by Novak and Cañas in research undertaken at Cornell University in the 1970s (Novak & Cañas, 2008; Novak & Musonda, 1991). Concept mapping is a visual representation tool to organise information about a chosen concept (Nesbit & Adesope, 2006). The concepts are often presented in enclosed shapes, and lines are used to connect any two concepts which are related. The most general concept that embraces all other concepts is presented at the top, and more specific concepts should be derived in descending order, hence reflecting a hierarchy of the presented concepts (Novak & Cañas, 2008; Schroeder et al., 2018). "Linking words or linking phrases" (Novak & Cañas, 2008, p. 1) are short descriptions on the lines that specify how the connected concepts are related. Some examples of such words/phrases include "is an example of", "generalises to" and "contains".

Novak (Novak & Cañas, 2008), the pioneer of concept maps describes the key features of concept mapping as the following:

- Hierarchy: The order in which the concepts are presented should imply a hierarchy from general concepts to specific concepts in a descending manner. However, the hierarchical relationship between concepts may differ with respect to the context. As a guide to deciding

the hierarchy between concepts, it is useful to have a "focus question" (Novak & Cañas, 2008, p. 2), a specific question that the concept mapper wishes to answer.
- Cross-links: Cross-links can be thought of as connections that are found between concepts that are derived from different strands developed from the main concept. The discovery of cross-links relies heavily on the creativity of the concept mapper.
- Examples: Having specific examples as a part of the concept map allows a clear understanding of a concept that may be abstract on its own. However, it is important to note that examples are not considered as concepts.

Benefits of concept mapping

Concept mapping is a great way for learners to make their internal understanding explicit (Croasdell et al., 2003; Kinchin et al., 2000). It provides an opportunity for the mappers to think in a critical and a complex (non-linear) way (Gul & Boman, 2006; Lee et al., 2013).

There are two major ways students can use concept mapping for their learning. It can either be given as a study guide completed by an expert for students to study with, or they can create a map independently. However, the Brod (2020)'s review showed that concept maps have a much greater impact when students create their own. This is because when a student creates their own concept map, the mapping process requires a meaningful engagement from the concept mapper, which involves higher-order learning activities such as organising and synthesising and, as a result, the task is very likely to enable high quality of learning (Nesbit & Adesope, 2006; O'Day & Karpicke, 2020; Schroeder et al., 2018). According to Moorf and Readence (1984), this may be a possible reason as to why concept mapping has a greater effect when done at the end of learning a topic rather than at the beginning (Moorf & Readence, 1984; Nesbit & Adesope, 2006).

The nature of concept mapping promotes the mapper's skill of showing their understanding in an organised way (Novak & Cañas, 2008; O'Day & Karpicke, 2020). Such skill is beneficial for "free recall" (O'Day & Karpicke, 2020, p. 2) of information (Hunt, 2012; Kahana, 2017; Raaijmakers & Shiffrin, 1981). Moreover, in the process of creating a concept map, students are expected to "select and isolate key pieces of information, organize that key information in a graphical form, and integrate those pieces of information together with relationship links" (Fiorella and Mayer (2015) as cited in O'Day and Karpicke (2020, p. 10). Such explanation is supported by the results of a recent experimental study conducted by O'Day and Karpicke (2020), which established a positive impact on learning when concept mapping and retrieval practice are incorporated into learning practice.

In certain contexts concept mapping has been shown to be more effective than conventional instructional practices (Jegede et al., 1990; Novak, 1990). Croasdell et al. (2003) have provided a specific comparison between learning through concept maps (abbreviated as CM) and linear note-taking (abbreviated as NT) to conclude that:
- CM is more effective for retrieving information compared to NT.
- Spotting important information is more easily done using CM than using NT.
- Relationships between concepts are easily noticeable in CM.
- Reviewing is less time-consuming and has a greater effect when using CM.
- The structure of CM is more flexible for adding new ideas.

However, when concept mapping is used as a form of assessment, there may be some disadvantages. Firstly, irrespective of the students' actual understanding of the chosen concept, the student may not have the required skill to make their understanding explicit or turn it into a given template. Secondly, because concept mapping is such a subjective activity that is highly dependent on the individual, consistency in the marking process cannot be guaranteed (McClure et al., 1999). In fact, the second point served as a foundation for the design, development and implementation of our study, which is reported in this paper.

Ways to evaluate concept mapping

Concept maps are not only a good tool for learning, but also for assessing students' understandings (Croasdell et al., 2003; McClure et al., 1999). However, as unique as concept maps can be, it is also very difficult for teachers to evaluate students' works.

Kinchin et al. (2000) define three different categories for considering the structural quality of concept maps. (Refer to Figure 1 below.) The first type (A in Figure 1) is "spoke" (Kinchin et al., 2000, p. 47), where the main concept is placed at the centre, and the only connections found in the map are between each derived concepts and the central concept, resembling the shape of a spoke. The second type (B in Figure 1) is "chain" (Kinchin et al., 2000, p. 47). A chain type of concept map shows a linear structure, developing only a single strand of a specific aspect of the main concept. Both "spoke" and "chain" are not considered good examples of concept maps. The last type (C in Figure 1), "net" (Kinchin et al., 2000, p. 47), shows a hierarchy by developing more specific concepts of each strand from the main concept and also reveals connections between concepts developed from different strands.

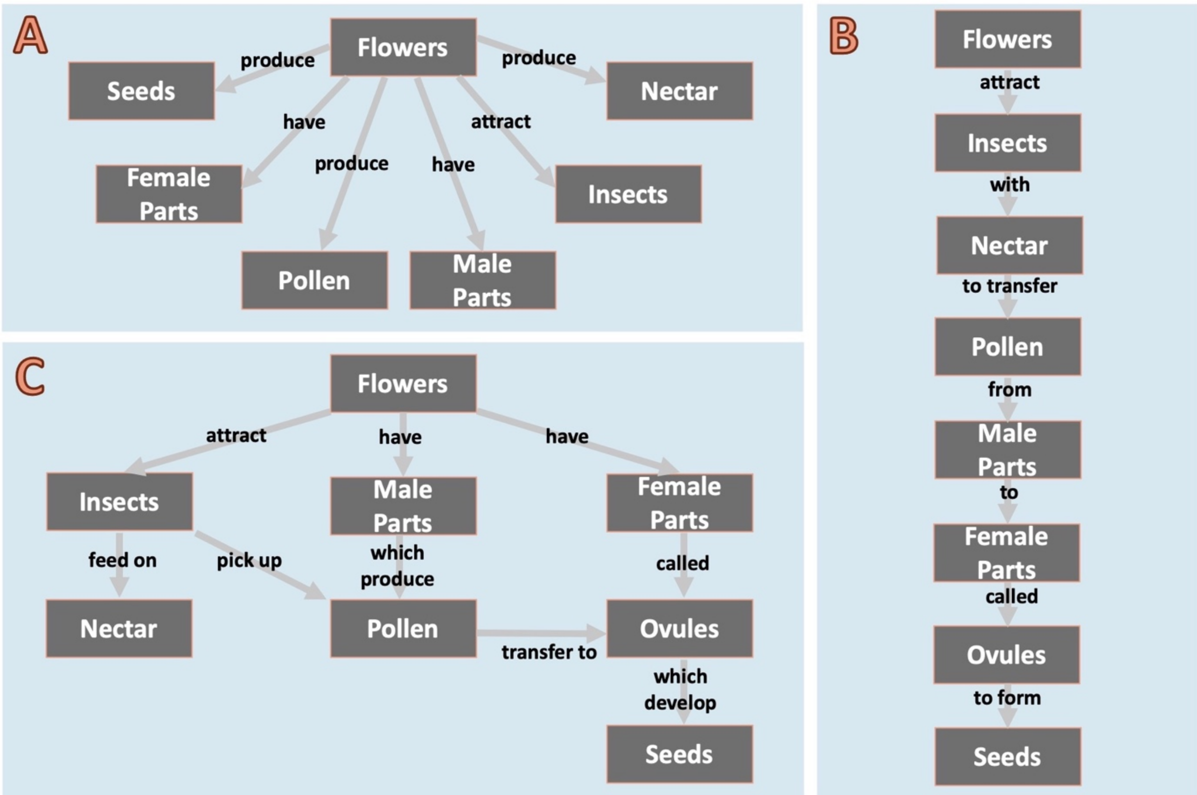


Figure 1: Structural types of concept maps, adapted from Kinchin et al. (2000)

Croasdell et al. (2003) provide very specific methods of evaluating concept maps. The components to be considered are:

- The total number of concepts identified;
- The total number of relationships found;
- The complexity of the map (subtract the minimum number of possible relationships when placed linearly from the total number of relationships used by the mapper);
- Comparison between the student's map and an expert's concept map;

- The progress made by the student during the course by comparing their concept mapping at different time points.

These components were explicitly defined by researchers of concept mapping for the purpose of research investigations to be used as an evaluation tool. However, no research provides a systematic way to evaluate learners' concept mapping activity that can be used in practice for mathematics university education. Addressing this gap, this study reports on the design, development, and implementation of concept mapping as part of a university mathematics course and reports on validation to identify an optimal method for assessing learners' concept maps.

Research questions

The main goal of this study was to find an optimal way that can be used in practice to evaluate concept maps. To achieve this goal, we devised a comparative analysis of students' scores when assigned according to the different rubrics. Specifically, we aimed to answer the following research questions:

- Which of the marking methods is the best predictor of learners' performance, measured as scores on the final exam?
- Which of the chosen marking methods provides the best way to assess student concept mapping activity?

METHODS

Research site

The study was conducted at a large research-intensive university (University of Auckland, New Zealand) in an undergraduate mathematics course covering Calculus II, Linear Algebra II, and Introduction to Ordinary Differential Equations, serving the needs of students majoring in a variety of disciplines. For a large proportion of non-mathematics majors taking this course, a lack of interest in the subject contributes to low intrinsic motivation, and suboptimal engagement with the course. An additional challenge is the size of the course: the enrolment numbers range from 350 to 550 students per semester. The course is delivered over 12 teaching weeks with the following weekly structure: three 1-hour lectures and one 1-hour tutorial (25 to 30 students per room working on problems).

This study was conducted in Semester 1, 2021 (March - June), when the COVID-19 pandemic has affected many places internationally. However, it should be noted that in New Zealand, due to the elimination strategy with closed borders, most educational institutions were functioning as normal from late 2020, with a few exceptions. Specifically, in Semester 1, 2021 at the University of Auckland, most courses were delivered face-to-face except for the first two weeks of the semester.

Participants

In the trial semester, 355 students were enrolled in the course, with 35 students studying overseas and completing the course online. An important component of this course is tutorials, which are practical sessions where students work on provided mathematics problems. All students are required to attend a tutorial each week for ten weeks. In addition, students are expected to submit their solutions to a "marked problem" weekly. A concept mapping task was given in each tutorial as a part of the question set and assigned twice as a marked problem.

Intervention: Knowledge Organisers and concept maps in tutorials

The intervention design was informed by the findings from experimental educational psychology pertaining to the learning-enhancing effect of the use of concept mapping in educational contexts. In each tutorial, students were expected to complete a Knowledge Organiser, starting from Week 2. Moreover, two concept mapping tasks were assigned as “marked problems”: in Week 3, students were asked to create a concept map on Series and in Week 7 on Vector Space. A template for Knowledge Organisers was provided (see Figure 2 on the next page). For a given concept, a Knowledge Organiser tasks students to state the definition of a given concept, provide at least two examples of the concept, state at least one non-example and create a concept map of the concept. The Knowledge Organisers were designed by the first author, with the help of the second author, who had been teaching this course for 15 semesters. The students were given a week to complete each Knowledge Organiser. Out of ten Knowledge Organisers that the students were expected to complete during the course, only two were collected for marking.

A short introductory session on Knowledge Organisers was provided in the first lecture by both authors, and an example of a Knowledge Organiser was uploaded on Canvas (Learning Management System) so that students could access it anytime. However, after receiving the first set of Knowledge Organisers submitted by students, the quality of the submissions revealed that some students did not understand the task clearly. Therefore, another explanation session was provided by a lecturer of the course (second author), during a lecture.

DATA COLLECTION

Data was collected from the learning management system (Canvas), providing marks for course assessment together with scores assigned by a researcher (first author) according to four different rubrics for evaluation of concept mapping activity.

Ethics approval was granted by the University of Auckland Human Participants Ethics Committee on 25/02/2021 for three years (reference Number UAHPEC21976).

Coursework marks were taken into account for this study. The course assessment structure comprised the following:

- 1 Final exam (50%)
- 1 Mid-semester test (20%)
- 30 Quizzes (15%)
- 10 Marked problems (10%)
- Tutorial participation (5%)

The final exam, the largest assessment component at the end of the semester, was held online. However, the mid-semester test was held on campus in-person (invigilated) for most students except for 35 students who could not enter the country due to the COVID-19 pandemic and had to take the test online. The mid-semester test was excluded from the data due to a significant difference in the performance of the two groups and concern that violation of academic integrity could translate into misleading results. Similarly, tutorial participation marks were excluded from the data due to substantial difference in getting credit for the two groups of students (face-to-face students gained marks for participating in tutorials, whereas online students had to submit written solutions).

The 30 quizzes were short online assessments that were due before each lecture from the second week of teaching. The lowest four marks were dropped. Hence, the data used in this study contains the top 26 quiz marks only.

Finally, the ten marked problems were short written assignments due weekly. Two of the marked problems were on concept maps, one on *Series* and the other on *Vector Space*. The guideline for Knowledge Organisers given to students is shown in Figure 3 below.

MARKED PROBLEM
Use the template below to create a Knowledge Organiser on the topic **vector spaces**.

Knowledge Organiser

Concept:
Definition:
Example:
Non−example:
Elaboration:

Figure 2: Knowledge Organiser template provided to students (Elaboration for concept mapping)

- **Concept**: Name the concept.
- **Definition**: Provide a definition of the concept.
- **Example**: Give two or more examples of the concept that are NOT in the course book.
- **Non-example**: Give at least one example of something similar but not the same as the concept given (NOT from the course book).
- **Elaboration**: Draw a diagram (concept or mind map) about the given concept using other concepts that are known to you, identifying the relations between them to organise and visualise the information.

Figure 3: Guidelines for completion of a Knowledge Organiser given in tutorials

A model solution was provided to the students with a disclaimer that there are many variations of a ‘correct’ concept map.

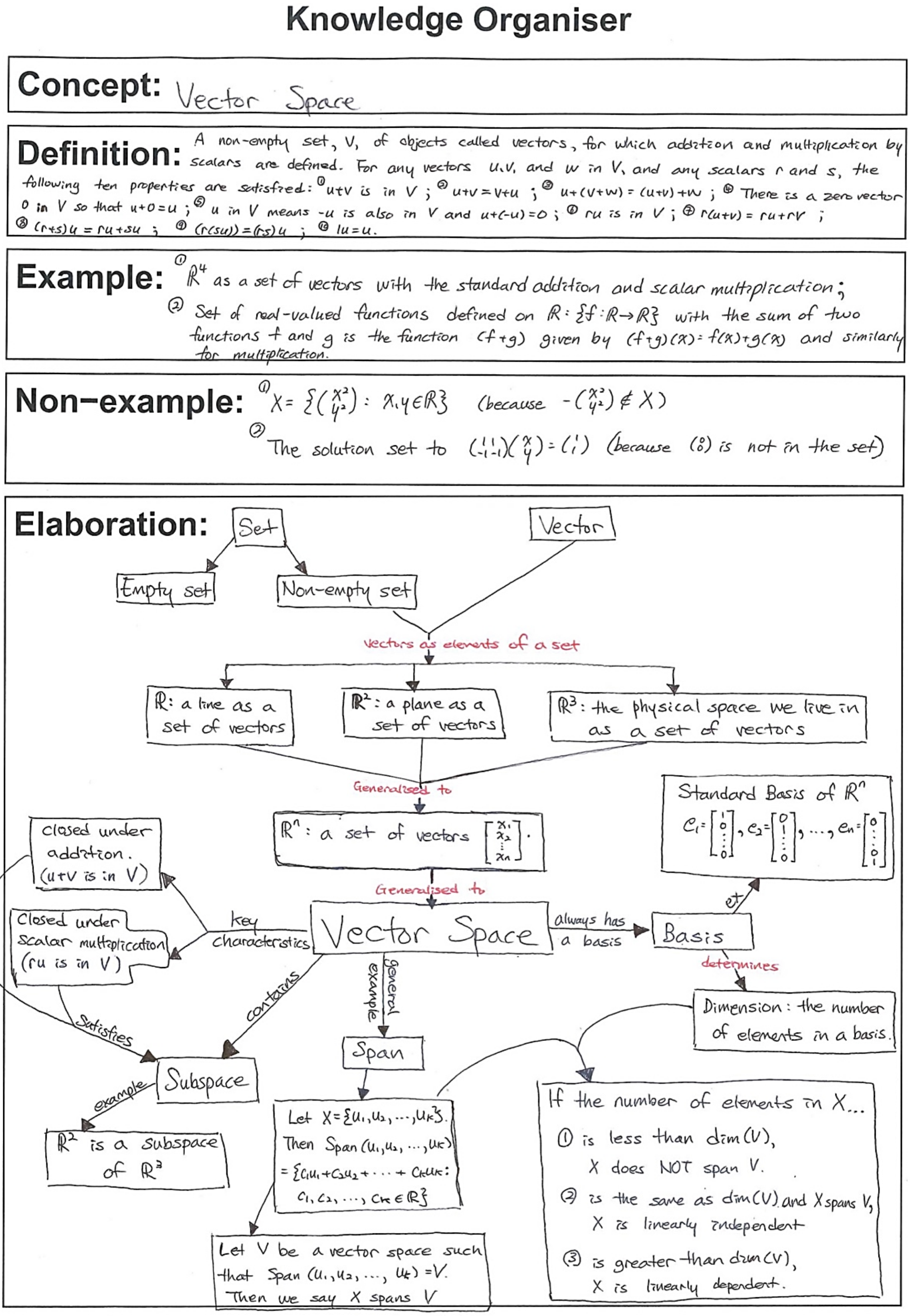


Figure 4: Model solution to a Knowledge Organiser (including concept mapping)

DATA ANALYSIS

As this research focused on finding an optimal way for analysing students' work in concept mapping, four different evaluation methods were selected. Specifically, the following methods suggested by other researchers were adopted with minor alterations to suit the practicalities of our educational context.

Method of Kinchin et al. (2000)

Firstly, the ideas of spoke, chain and net by Kinchin et al. (2000) were manipulated. The maximum marks allowed were 3. It was suggested by Kinchin et al. (2000) to award marks as follows:

- 0/3: Nothing was provided, or the work produced is not in the form of a concept map.
- 1/3: A chain or spoke type of concept map was produced.
- 2/3: A net type of concept map was produced with weakly developed connections.
- 3/3: A net type of concept map was produced with highly developed connections.

However, considering that the students may not have experienced such tasks before, we were concerned that a highly strict marking scheme may cause students to lose motivation. Hence, including "effort" into consideration, a modified marking scheme was produced, which was used in the course:

- 0/3: Nothing was provided, or the work produced is not in the form of a concept map.
- 1/3: A chain or spoke type of concept map was produced with weak evidence of effort.
- 2/3: A chain or spoke type of concept map was produced with strong evidence of effort. OR A net type of concept map was produced with weak evidence of effort
- 3/3: A net type of concept map was produced with sufficient evidence of effort.

Methods of Croasdell et al. (2003)

As previously mentioned, Croasdell et al. (2003) proposed many different aspects to consider in evaluating the quality of concept maps. Out of the five components listed in the Introduction section, only the first three (the number of concepts, the number of relationships, and the map's complexity) were chosen. The complexity is measured by a numeric value, which is equal to the inverse ratio of the number of concepts and the number of relationships between them.

We did not include the last two aspects outlined by Croasdell et al. (2003) (listed in the Introduction section) because of the following reasons; first, comparing the students' maps to a map of an expert seemed to be contradicting the fact that the concept mapping task should be a creative and idiosyncratic process. Also, the last aspect (tracking development through different time points of the course) was not suitable for our study because the concept maps were collected for marking only twice during the course.

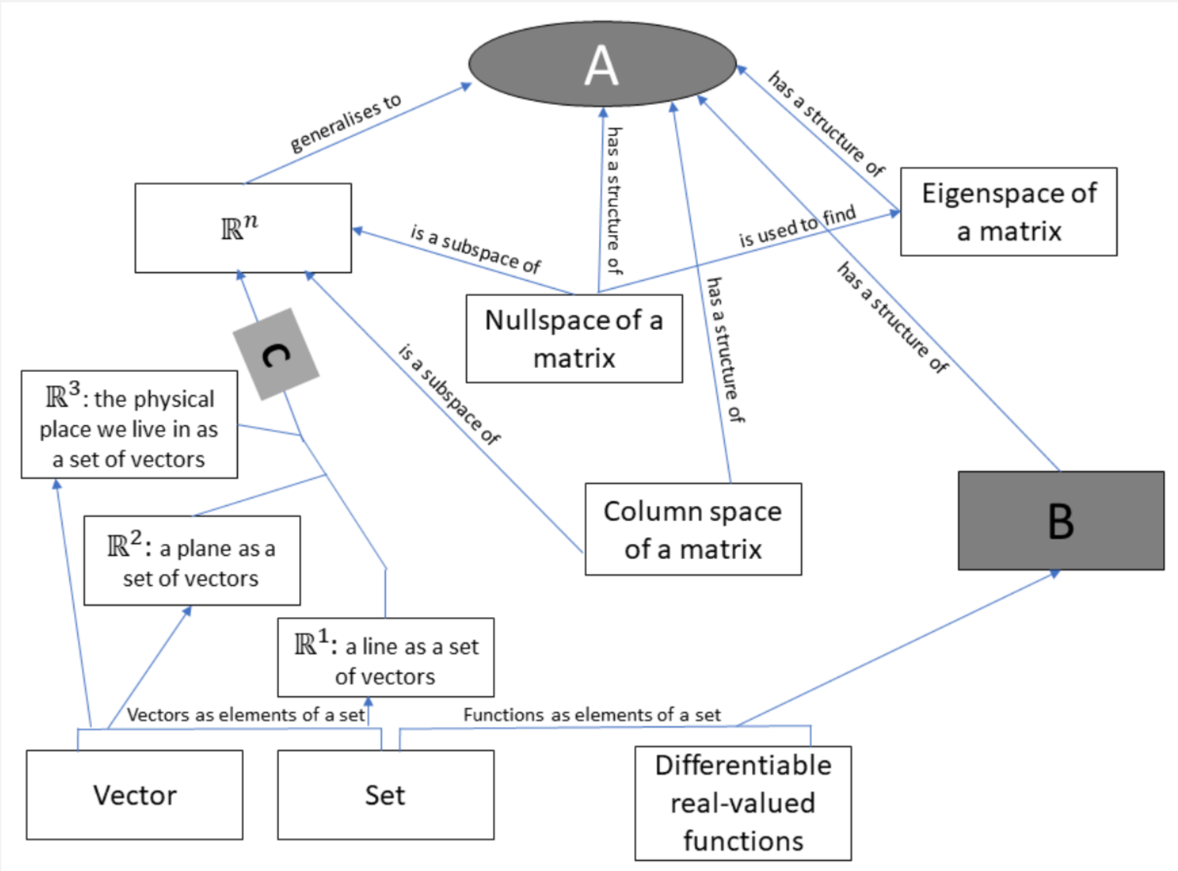
The four rubrics used in the study: Structure, Concept Count, Relationship Count and Ratio methods

For convenience, the methods introduced in this section will be replaced by concise descriptions to label the four rubrics compared in this study. The modified method of Kinchin et al. (2000) will be described as the *Structure* rubric. Out of the methods introduced by Croasdell et al. (2003), the evaluation method with a focus on the number of concepts will be referred to as the *Concept Count* rubric, and the method counting the number of relationships will be labelled as the *Relationship Count* rubric. Lastly, the method of analysing the complexity as the inverse ratio between the concepts and the relationships used in concept mapping will be described as the *Ratio* rubric.

Exam concept map question

The final exam comprised 30 multi-choice questions, with one question focusing on concept mapping. The concept map question had a form of a completed concept map with a few missing concepts and relationships for students to figure out. Here is the question:

This question has three parts - make sure you answer all of them.



Consider this incomplete concept map. Fill out the three blanks: The main concept **A** in the grey ellipse is

Vector space . The concept missing in grey box **B** is

Set of solutions of a linear differential equation . The missing connecting relation labelled by **C** is

generalises to

Figure 5: Concept map question in the final exam

Students were presented with three *in-line choice* questions as part of the question with multi-choice options to replace objects A, B and C in the concept map. The multi-choice options given for each of A, B, and C are shown in Figure 6, with the correct answers at the bottom of the lists:

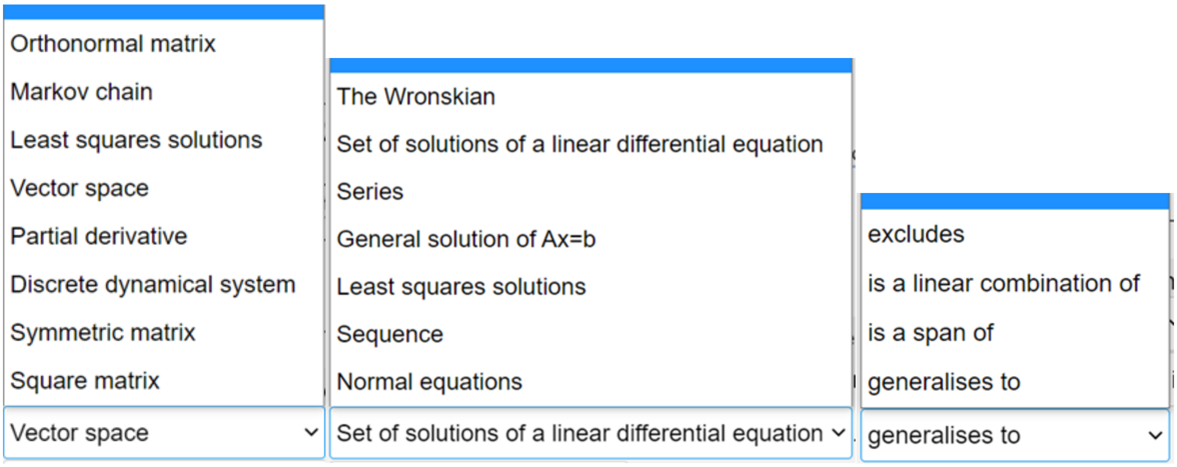


Figure 6: Options for A, B and C

Considering that it may be the first time that students would have encountered such a question in an exam context, a mock exam was provided containing a similar question, allowing for formative practice.

RESULTS AND DISCUSSION

This section compares the four rubrics for the evaluation of concept maps produced by students as part of the coursework assessment by comparing multiple regression models to predict their final exam outcomes. The final exam scores were chosen for this analysis because by constructing one's own concept map, a meaningful engagement will be required from the concept mapper. From research, it is known that concept mapping involves higher-order learning activities such as organising and synthesising and, as a result, it is expected to enable high quality of learning (Nesbit & Adesope, 2006; O'Day & Karpicke, 2020; Schroeder et al., 2018).

The main goal of this analysis was to identify whether the students' performance of the concept mapping task was an appropriate predictor of their overall performance. Thus, the final exam scores were chosen as a measure of their overall performance with separate consideration of student performance on the concept map exam question.

In order to identify the most appropriate marking rubric, a multiple regression was done using scores from different marking rubrics, the total quiz score and the total marked problem scores (excluding the two concept map problems) as the independent variables. The final exam outcomes were selected as the dependent variable. Specifically, two dependent variables were considered separately: the score on the concept map question only and the total final exam score.

Predicting Exam Concept Map Question Score

Table 1 shows the results of the first four tests with the score on the concept map exam question as the dependent variable. After data cleaning, four multiple regressions were run to predict the exam concept map question score from the concept map scores using each evaluation method, the total quiz score and the total marked problem score. The assumptions of linearity, independence of residuals and homoscedasticity were satisfied. There was no evidence of multicollinearity. There was one unusual point, identified by its studentised deleted residuals. However, no problem was found with the values of the data point. Hence, there was no reason to delete the data point. The assumption of normality was met.

Table 1: Multiple regression results for Exam concept map question score (Models 1-4)

Exam concept map question score	*B*	95% CI for *B* LL	95% CI for *B* UL	SE *B*	β	R^2	ΔR^2
Model 1						.044	.032*
Constant	1.642***	.736	2.547	.460			
Structure method	.100*	.023	.177	.039	.161*		
Total quiz score	-.001	-.013	.011	.006	-.008		
Total marked problem score	.010	-.005	.025	.007	.112		
Model 2						.030	.018
Constant	1.678***	.765	2.591	.464			
Concept count method	.012	-.002	.027	.007	.106		
Total quiz score	.000	-.013	.012	.006	-.006		
Total marked problem score	.011	-.003	.026	.007	.124		
Model 3						.037	.026*
Constant	1.693***	.783	2.602	.462			
Relationship count method	.015*	.002	.028	.007	.139*		
Total quiz score	-.001	-.013	.011	.006	-.009		
Total marked problem score	.011	-.004	.025	.007	.119		
Model 4						.045	.034**
Constant	1.623***	.718	2.528	.459			
Ratio method	.264**	.067	.461	.100	.164**		
Total quiz score	-.001	-.013	.012	.006	-.007		
Total marked problem score	.010	-.004	.025	.007	.117		

Note. Model = "Enter" method is SPSS Statistics; *B* = unstandardized regression coefficient; CI = confidence interval; LL = lower limit; UL = upper limit; SE *B* = standard error of the coefficient; β = standardized coefficient; R^2 = coefficient of determination; ΔR^2 = adjusted R^2. *$p < .05$. **$p < .01$. ***$p < .001$.

Models 1, 3 and 4 were statistically significant in predicting the exam concept map question score, but Model 2 was not. (Model 1: $F(3, 256) = 3.890$, $p = .010$; Model 2: $F(3, 256) = 2.617$, $p = .052$; Model 3: $F(3, 256) = 3.321$, $p = .020$; Model 4: $F(3, 256) = 4.010$, $p = .008$)

In each of models 1,3, and 4, only one of the variables added statistically significantly to the prediction. Those variables were the concept map scores using the *Structure* method ($p = .011$), the *Relationship count* method ($p = .027$) and the *Ratio* method ($p = .009$), respectively. In Model 2, none of the variables added statistically significantly to the prediction.

In conclusion, Model 4 has the best fit with the highest adj. R^2 value, and with the highest value of the unstandardised regression coefficient (B=0.264), thus suggesting that the *Ratio* method of evaluating student concept mapping activity is the most accurate.

Predicting Exam Score

Table 2 shows the results of the four multiple regression tests with the total exam score as the dependant variable, using each of the four evaluation rubrics, the total quiz score and the total marked problem score as independent variables. The assumptions of linearity, independence of residuals and homoscedasticity were satisfied. There was no evidence of multicollinearity. There were no unusual points, and the assumption of normality was met.

Table 2: Multiple regression results for Total Exam score (Models 5-9)

Total exam score	*B*	95% CI for *B*		SE *B*	β	R^2	ΔR^2
		LL	UL				
Model 5						.192	.183***
Constant	6.730***	1.398	12.061	2.707			
Structure method	.560*	.108	1.013	.230	.140*		
Total quiz score	.060	-.011	.131	.036	.126		
Total marked problem score	.167***	.081	.252	.043	.292***		
Model 6						.191	.182***
Constant	7.025*	1.685	12.366	2.712			
Concept count method	.102*	.017	.187	.043	.136*		
Total quiz score	.057	-.014	.129	.036	.120		
Total marked problem score	.172***	.087	.257	.043	.300***		
Model 7						.195	.186***
Constant	7.084**	1.757	12.411	2.705			
Relationship count method	.103**	.026	.179	.039	.152**		
Total quiz score	.057	-.014	.128	.036	.120		
Total marked problem score	.169***	.084	.255	.043	.296***		
Model 8						.201	.192***
Constant	6.604*	1.300	11.907	2.693			
Ratio method	1.738**	.583	2.892	.586	.168**		
Total quiz score	.059	-.012	.130	.036	.124		
Total marked problem score	.168***	.083	.253	.043	.294***		

Note. Model = “Enter” method is SPSS Statistics; *B* = unstandardized regression coefficient; CI = confidence interval; LL = lower limit; UL = upper limit; SE *B* = standard error of the coefficient; β = standardized coefficient; R^2 = coefficient of determination; ΔR^2 = adjusted R^2. *$p < .05$. **$p < .01$. ***$p < .001$.

All of the models 5, 6, 7 and 8 were statistically significant in predicting the total exam score (Model 5: $F(3, 256) = 20.311$, $p < .001$; Model 6: $F(3, 256) = 20.171$, $p < .001$; Model 7: , $F(3, 256) = 20.734$, $p < .001$; Model 8: $F(3, 256) = 21.455$, $p < .001$).

In models 5, 6, 7, and 8, two of the variables added statistically significantly to the prediction. The total marked problem score was significant in all four models ($p < .001$ in all four models). The other significant variable in models 5, 6, 7, and 8 were the concept map scores using the

Structure method (p = .015), the *Concept count* method (p = .019), the *Relationship count* method (p = .009) and the *Ratio* method (p = .003), respectively

Out of the four models, Model 8 reported the highest unstandardised regression coefficient of the evaluation method (Model 5: B=.560; Model 6: B=.102; Model 7: B=.103; Model 8: B=1.738) and the highest adj. R^2, thus suggesting that the *Ratio* method is better suited for marking student concept mapping tasks.

Discussion

Out of the eight tests that were conducted, all models except Model 2 were statistically significant. In models 1, 3 and 4, the concept map scores obtained from different marking rubrics turned out to be the only significant independent variable. In models 5, 6, 7 and 8, the concept map scores from different marking rubrics as well as the total marked problem score were significant independent variables. This is an expected result as the students' concept mapping performance is most likely to be directly related to their ability to recognise correct concepts and relations in the given concept map on the exam. On the other hand, the total marked problem score consists of marks from mathematical problems on different topics. Hence, it is very likely to be a significant predictor of the overall performance on the final exam.

The total quiz score was not a statistically significant predictor in any of the tests. A possible reason for this is that the quizzes were designed to provide an impetus for revision after every lecture so that the frequency of student engagement is increased in order to improve student self-efficacy (Evans et al., 2021; Riegel & Evans, 2021). The questions in the quizzes are academically less demanding compared to creating a concept map or completing a marked problem. Moreover, students were allowed 2 attempts at each quiz, which gave them a higher chance of scoring full marks and the lowest four scores were dropped. Therefore, it is less likely to reflect students' understanding of the mathematical content accurately.

According to the results of these models, the concept map scores from the *Structure* and the *Ratio* rubrics seem to be the best predictors of the students' performance on the exam concept map question. The models using these two marking rubrics showed the highest adjusted R^2 values (Model 1: adj. R^2 = 0.032; Model 4: adj. R^2 = 0.034). Comparing the effects of the different marking rubrics as a predictor of the final exam performance (total), we observed that the *Ratio* method showed the highest adjusted R^2 value of 0.193, with the highest B value (B=1.674).

This shows that the rubric utilising the ratio of the number of concepts and the number of relationships best reflects students' overall mathematical performance. Hence, the most important factor to consider when evaluating students' concept maps is the proportion and not just the count of concepts and connections that a learner comes up with.

Then, why might the *Structure* method scoring also be a significant predictor of the exam concept map question with the second-best indicators? Unlike the concept mapping tasks that the students were given as a marked problem in the tutorials, the exam concept map question required the students to comprehend a concept map that was already nearly completed by another person. Therefore, in this process, it was necessary for students to comprehend the structure of the concept map. Hence, the scores from the *Structure* method were likely to factor heavily in predicting the students' performance in this particular exam question.

CONCLUSION

In this study, we demonstrated how concept mapping activities can be incorporated in a mathematics course and found an optimal method to evaluate concept maps created by students, thus providing recommendations for implementation in practice. Out of the methods tested in this study, the *Ratio* method was identified to be the most comprehensive way of evaluating students' concept mapping. However, in predicting performance on a ready-made concept map on a final exam, two methods appear adequate for assessment: the *Structure* method as well as the *Ratio* method.

The structure method's advantage is that it is not just a formulaic way of evaluating as it captures the various outputs of the concept maps made by the students, such as examples and explanations. However, this method may result in inconsistency in the marks given out since the marking process may be subjective

Taking this consideration into account, we conclude that the *Ratio* method is the most optimal evaluation method. As Novak, the creator of concept maps, outlines in his study, cross-link is one of the main features of concept maps (Novak & Cañas, 2008). This shows that only having many concepts with no cross-links overlooks the major aims of creating concept maps. On the other hand, because the cross-links connect two concepts, it is impossible to have a "relationship" that stands alone without linking any concepts. That is, relationships cannot exist without the presence of concepts. Hence, the *Ratio* method, which considers both features is most likely to evaluate the concept maps accurately. Most importantly, an optimal concept map focuses on finding maximal relationships within the selected concepts. In particular, if there were two concept maps with the same number of concepts, the map with more relationships identified is considered to be of higher quality. In this sense, the *Ratio* method perfectly captures this key feature of concept mapping. However, there is a caveat to consider when a concept map is minimalistic yet has a high inverse ratio of concepts and relations. For example, a concept map with two or three concepts only. In this case, we recommend including evaluation of effort as part of the marking rubric so that students do not use this loophole.

In summary, this study demonstrated successful incorporation of concept mapping activities as part of a large undergraduate mathematics course. It showed promising results pertaining to the utility of concept mapping as a learning-enhancing tool. Students' performance on concept mapping task was a significant predictor of their exam scores. This suggests that further implementation studies could be conducted with a focus (1) on establishing causal relationships between the use of concept mapping and learning outcomes, (2) investigating to what extent an increase in concept mapping activity throughout the semester would enhance student conceptual understanding, (3) researching different ways to incorporate concept mapping activity in a course, for example, as a group task in face-to-face tutorials, enabling collaborative interactions.

REFERENCES

Brod, G. (2020). Generative Learning: Which Strategies for What Age? *Educational Psychology Review*, 1-24.

Croasdell, D. T., Freeman, L. A., & Urbaczewski, A. (2003). Concept maps for teaching and assessment. *Communications of the Association for Information Systems*, *12*(1), 24.

Evans, T., Kensington-Miller, B., & Novak, J. (2021). Effectiveness, efficiency, engagement: Mapping the impact of pre-lecture quizzes on educational exchange. *Australasian Journal of Educational Technology*, 163-177. https://doi.org/10.14742/ajet.6258

Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity*. Cambridge University Press.

Gul, R. B., & Boman, J. A. (2006). Concept mapping: A strategy for teaching and evaluation in nursing education. *Nurse education in practice*, *6*(4), 199-206.

Hunt, R. R. (2012). Distinctive processing: The co-action of similarity and difference in memory. In *Psychology of learning and motivation* (Vol. 56, pp. 1-46). Elsevier.

Jegede, O. J., Alaiyemola, F. F., & Okebukola, P. A. (1990). The effect of concept mapping on students' anxiety and achievement in biology. *Journal of research in science teaching*, *27*(10), 951-960.

Kahana, M. J. (2017). Memory search. *Learning and memory: A comprehensive reference*, 181-200.

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.

Lee, W., Chiang, C.-H., Liao, I.-C., Lee, M.-L., 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.

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.

Moorf, D. W., & Readence, J. F. (1984). A quantitative and qualitative review of graphic organizer research. *The Journal of Educational Research*, *78*(1), 11-17.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of educational research*, *76*(3), 413-448.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of research in science teaching*, *27*(10), 937-949.

Novak, J. D., & Cañas, A. J. (2008). The theory underlying concept maps and how to construct them. *Florida Institute for Human and Machine Cognition*, *1*.

Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, *28*(1), 117-153.

O'Day, G. M., & Karpicke, J. D. (2020). Comparing and combining retrieval practice and concept mapping. *Journal of Educational Psychology*.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological review*, *88*(2), 93.

Riegel, K., & Evans, T. (2021). Student achievement emotions: Examining the role of frequent online assessment. *Australasian Journal of Educational Technology*, 75-87. https://doi.org/10.14742/ajet.6516

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. In: Springer.