Research Article

International Journal of Advanced Technology and Engineering Exploration, Vol 12(125)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2024.111100266

Graphing knowledge: automated concept map evaluation to assess and enhance student learning

Sunu Mary Abraham[1, 2*] and G. Sudhamathy[1]
Department of Computer Science, Avinashilingam Institute for Home Science, and Higher Education for Women, Coimbatore, India[1]
Department of Computer Science, Rajagiri College of Social Sciences (Autonomous), Kerala, India[2]

Received: 11-March-2024; Revised: 10-April-2025; Accepted: 18-April-2025

Abstract
A concept map (CM) is a graphical representation of relationships and interconnections between concepts or ideas, using nodes and links to depict the hierarchical and interconnected structure of knowledge. A literature review on methods for evaluating students' CMs was conducted, revealing potential areas for further research in CM evaluation techniques. To address these gaps, this study introduces a novel framework, graphing students' knowledge using concept maps (GSKCM). GSKCM employs a graph-based scoring algorithm to automate the evaluation of concept maps and assess students' structural comprehension, thereby facilitating targeted remediation. The algorithm compares student-generated CMs to a reference map, validating individual concepts and propositions while emphasizing the hierarchical structure—ultimately assessing students' understanding of the topic's overall framework. The GSKCM framework was implemented in an authentic classroom setting, where students constructed CMs on a specific topic. These maps were evaluated using the framework, revealing both strong and weak conceptual areas for individual students and the class as a whole. Based on the algorithm's results, students were grouped into five clusters, enabling tailored tutoring at both the individual and group levels. An experimental study comparing the GSKCM approach to traditional evaluation methods demonstrated that students in the experimental group outperformed those in the control group on a formative assessment. By automating the evaluation process, GSKCM streamlines assessments, provides deeper insights into student understanding, and enhances the learning experience—particularly in mastering complex subjects.

Keywords
Concept maps, Automated assessment, Graph-based evaluation, Student knowledge representation, Educational technology, Personalized learning.

1.Introduction
A concept map (CM) is a graphical representation that illustrates the relationships between concepts using labeled links or connecting phrases. In a CM, a link label—also known as a relationship—is the word or phrase written on the line connecting two nodes, indicating the nature of their connection[1]. For example, in *Figure 1*, the phrases "provide," "is used to make," and "required by" serve as link labels connecting various concepts. The process of understanding these interconnections is known as concept mapping. A proposition in a CM is a meaningful statement formed by linking two or more concepts with appropriate connecting words or phrases.

For instance, the sentence *"Trees → provide → Oxygen → required by → living organisms"* represents a proposition derived from a CM. This hierarchical structure, with root, internal, and leaf nodes, helps in visualizing how knowledge is organized from general to specific, providing a clear representation of the relationships and dependencies between different concepts [1]. A CM is a straightforward and efficient method for assessing student knowledge, misconceptions, and learning retention [2]. Not only do misunderstandings negatively impact students' learning, but they also impede the development of their reasoning skills, both of which are important considerations. CMs offer an orderly grouping and organisation of knowledge, helping students grasp the relationships between concepts and enabling them to solve problems more easily [3, 4]. In education, CMs are

*Author for correspondence

frequently used to identify and resolve misunderstandings. Novak & Gowin introduced the CM technique, which was supported by Ausubel's meaningful learning theory[5]. Hierarchical structures [6] play a significant role in learning complex concepts through CM in evaluating the depth of understanding, particularly in science education.

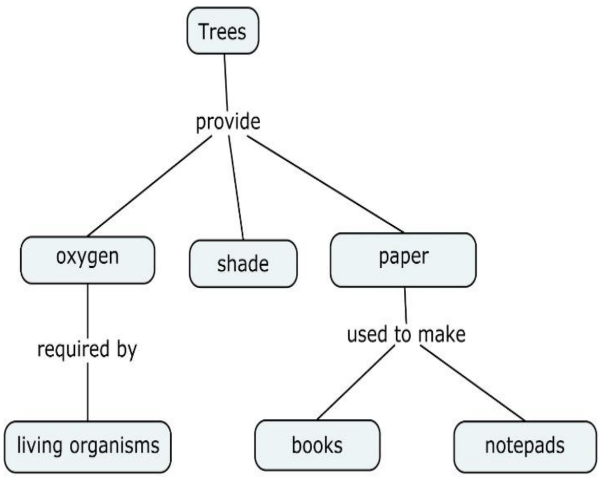


Figure 1 Example for a hierarchical concept map

The evaluation of CMs has been a subject of extensive research. Traditional assessment methods involved qualitative analysis, where educators manually compared student-generated maps with expert maps. However, this method is labour-intensive and subjective, leading to the need for the development of more structured evaluation techniques. Evaluating CMs presents several limitations and challenges that can impact the accuracy and effectiveness of the assessment process. One of the primary challenges in evaluating CMs is the inherent subjectivity, especially in holistic and propositional scoring [7–9]. It is difficult to achieve reliable results with holistic scoring, particularly in online learning environments where there may be less interaction between students and evaluators, which can clarify misunderstandings. Manual scoring methods, whether structural, propositional, or holistic - can be extremely time-consuming, especially in large educational settings.

Automating the evaluation of CMs helps instructors assess their students' understanding of a topic and take remedial actions more efficiently [10]. The objective of the research on the graphing students' knowledge using concept maps (GSKCM) framework is to develop an automated scoring system that evaluates the structural hierarchy and conceptual relationships within student-generated CMs. Additionally, students are clustered based on their scores, enabling personalized remediation strategies tailored to each student's level of comprehension.

This research makes a significant contribution by introducing the innovative GSKCM framework, which automates the evaluation of CMs using a novel graph-based scoring algorithm. This approach not only improves efficiency and consistency in assessment but also rigorously validates the structural hierarchy of concepts within the maps. By focusing on both the accurate representation of concepts and the validation of their hierarchical relationships, the framework enables a more comprehensive assessment of students' structural understanding, leading to tailored remediation and improved educational outcomes.

The GSKCM framework evaluates a student's CM against a reference CM and generates targeted recommendations. The scoring algorithm assigns scores to concepts at different hierarchical levels, allowing the identification of weaker concepts, unknown concepts, and misconceptions. These insights guide students in reconstructing their understanding of the topic, fostering deeper and more accurate knowledge acquisition.

The research paper is organized as follows: Section 2 presents a comprehensive review of existing literature, comparing various scoring algorithms and evaluating the limitations of current approaches. Section 3 details the methodology and system framework, including an in-depth explanation of the proposed scoring algorithm. In Section 4, the framework is applied to evaluate students' CMs in an educational setting, followed by an analysis of the results. Section 5 compares the proposed framework with alternative CM evaluation methods. Finally, conclusions are drawn in Section 6.

2.Literature review

CMs are visual methods for organising and displaying knowledge that have been utilised in various areas of education and training. Construction of CMs by students facilitates meaningful and effective learning. The knowledge structure of students' CMs can be evaluated with or without a provided list of concepts [11]. Without the list of concepts, a simplified technical concept structure is created with fewer cross- linkages. Students analyze each concept and its role, building a tight knowledge structure using only the concepts from the list.

An extended scratch-build and an extended kit-build on students' knowledge structures and perceptions using CMs are compared in [12]. Fifty-five second-year university students participated in both the control and experimental groups. The control group used the extended scratch-build map, and the experimental group used the extended kit-build idea mapping tool. Students' knowledge structures were measured by proposition quality and structural map scores. This study used the technology acceptance model to confirm students' concept mapping tool perceptions. The experimental group demonstrated significantly higher proposal quality and structural map ratings than the control group. This study also indicated that extended kit-build outperformed the extended scratch-build in terms of perceptions. A limitation is that the size of CM components may not consistently reflect the quality of the knowledge structure.

Text analysis and association rule mining [13] are used to generate CMs automatically. The algorithm automatically and quickly generates the CM, eliminating the need of experts for evaluation, and dynamically updates it based on criteria like test question confidence thresholds. The main drawback is that it fails to accurately capture the nuances of complex concepts and their relationships, as it relies solely on text analysis and association rule mining. Jain et al. [14] presented an artificial intelligence (AI)-based student learning evaluation model. This method uses CMs to evaluate students' subject knowledge. The scoring method generates probability distribution curve-fitting scores for each student's CM. For CM evaluation, the probability distribution curve of scores for the students' concepts was analysed. Positive findings were obtained from two undergraduate classes that used the program. A limitation of the method is that the instructor manually scores every CM. Therefore, instructors must choose the best model for their purposes and course scope. Gurupur et al. [15] discuss how CMs and Markov chains can be used to gauge a student's grasp of a subject. The CM identifies three levels of detail: Gist level, Support level, Detail level. A finite Markov chain with 3 states is identified to analyse the transition from one state to another. This method effectively resolves issues involving cyclic chains of concepts. Although Markov chains can be used to assess CMs with circular chains of concepts, the validation of concepts is still performed manually. A new fuzzy based inference system model for concept maps (FISCMAP) [16] was proposed that integrated fuzzy logic constructs with concept mapping. Here, eight CM variables were extracted and used as inputs to a five-level fuzzy inference system equipped with fuzzy rules. Students and academic institutions are given feedback on the quality of CM as a reflection to guide them toward an intelligent and enriched learning environment. One limitation is that the CM is evaluated holistically, without individual concept scores. Another issue is the reliance on predefined fuzzy rules, which might not accommodate the diverse learning styles and dynamic student needs in real-time.

A new validation approach for CMs is suggested in [17], where the validity of concept is divided into content validity and application validity. A CM on geometry is compared with empirically collected criterion maps using similarity methods to determine its content validity. Knowledge space theory is used to predict problem-solving behaviour, thereby establishing application validity. However, the validation methodology may not account for all contextual factors that influence the applicability of CMs in diverse educational settings. Shou et al. [18] proposed an algorithm to discover learning behaviour patterns and organise the knowledge points. The approach developed a knowledge point interaction - degree model and a directed learning path network. Kullback-Leibler divergence (KLD) based-matrix learning behaviour similarity analysis using vectors of eigenvalues is employed here. A clustering algorithm analyzes the optimal learning path. An unsupervised feature selection (UFS) machine learning (ML) model [19] is employed using eight UFS methods to score twelve CM assessment features. The study analyzes` 202 maps from four case studies to identify informative and uninformative aspects and to observe context-related nuances. The number of edges may be less meaningful compared to the average degree, which is evaluated less frequently. Map diameter proves useful when learners construct maps based on detailed research but is less informative when maps are elicited based on learners' personal views. Future work may involve eliciting instructors' grading preferences and integrating them into an algorithmic approach (i.e., UFS) to generate a preferred feature ranking. The eight UFS algorithms offer five methods for ranking features within the maps. A noted limitation is that selecting informative features alone may not fully capture the complexity of student learning. Li et al. [20] provided an adaptive learning system using a learning path generation (LPG) algorithm. The students were clustered based on their mastery of the topic, and for each cluster, association rules mining was used to generate CMs, and topology sorting to generate learning paths for each cluster of students in the LPG algorithm. LPG customized learning paths to student achievement. A network analysis [21] of the CMs created by students for introductory psychology revealed that CMs differed between students, and network structure significantly predicted quiz scores. This research shows how network science can quantify learners' conceptual structures. Kroeze et al. [22] presented a tool that automatically evaluated student CMs and provided feedback based on a reference CM. This application can evaluate CM and provide accurate and useful feedback on a number of common flaws in students' CMs. Using word embeddings, another research [23] proposed a new graph-based CM feature for evaluating Chinese composition concepts. The notions and co-occurrence of the essay's concepts are depicted in the students' CM. Using eighth-grade student compositions from a large-scale test, the computer evaluation method achieved a higher score than the benchmark. The results demonstrate that the proposed strategy increases construct-relevant coverage in automatic concept evaluation of compositions and provides students with constructive feedback. A study by [24] evaluates and analyzes the use of open-ended CMs in knowledge and learning research. It explores analysis methods, CM properties, and their relationship to knowledge and comprehension outcomes. The review includes twenty-five studies involving open-ended idea maps. It examines the connections among the three research components and proposes criteria for analyzing methodological coherence. The empirical study [25] evaluated how CM quality affects feedback providers' knowledge acquisition. It also examined how pre-existing knowledge influenced this process. Secondary-school students were assigned to one of the three situations with varying CM quality levels to provide feedback while utilizing an online physics inquiry learning platform. An analysis examined whether the condition affected feedback providers' learning. This analysis examined post-test knowledge scores, students' CMs, and feedback. After providing feedback on lower-quality CMs, students scored higher on the post-test. Students who participated in a teacher-guided discussion after working with an expert CM and reflection prompts achieved the greatest learning gains, largely due to the higher quality of their reflections [26]. This underscores the effectiveness of guided discussions in enhancing learning outcomes (LOs) while using CMs. The CMs were found to be as effective as open-ended exams in assessing nursing students' knowledge, offering higher grading efficiency and more reliable gap identification [27]. A meta-analysis of 55 studies [28] found that CMs moderately enhance science achievement, with a stronger impact in physics and earth science compared to biology and chemistry. Learning improvements were influenced by teaching conditions, student demographics, and publication-related factors. Veiga et al. [29] analyzed the effectiveness of CMs in evaluating LOs in engineering courses, using data from individual and group maps across various academic levels. The findings suggest that CMs, evaluated through rubrics, self-assessments, and final exams, provide a suitable alternative for assessing understanding in online teaching situations.

Waterloo analysis [30] uses a qualitative rubric to grade CMs based on structural and temporal aspects. Topological scoring [31] evaluates map quality using automated structural metrics like concept and proposition counts. Fuzzy-based techniques [32] employ a fuzzy inference system for holistic, automated scoring, though they do not assess individual concepts.

The reviewed literature highlights the advantages of CMs in educational settings and presents various frameworks for evaluating CMs, focusing on their effectiveness in both learning and assessment processes [33–35]. However, a common limitation of these tools is the need for manual validation of concepts by instructors. Although some frameworks incorporate automated validation, they typically limit this to proposition checking and verifying the correctness of concept keywords. Additionally, many studies employ network analysis to compare and evaluate CMs using different similarity metrics [36–38]. Despite these efforts, such validations often overlook the hierarchical structure of concepts—an essential element for assessing a student's structural understanding of the subject.

3.Materials and methods

The web-based framework GSKCM depicted in *Figure 2*, developed in hypertext pre-processor (PHP) enables the student to quickly create CMs based on their subject-matter expertise to enhance the learning results. The student CM and the reference CM are stored in a MySQL database. The list of concept keywords provided by the instructor is used to generate CMs. Students develop CMs based on the material they have learned in order to visualize the topic's concepts and reorganise their connections. After construction, the maps are saved in the database and are compared with the reference CM using the scoring algorithm proposed here. By obtaining scores, students can assess their understanding of each concept. These evaluations can assess how well a subject or topic has been learned by an individual student or a group of students.

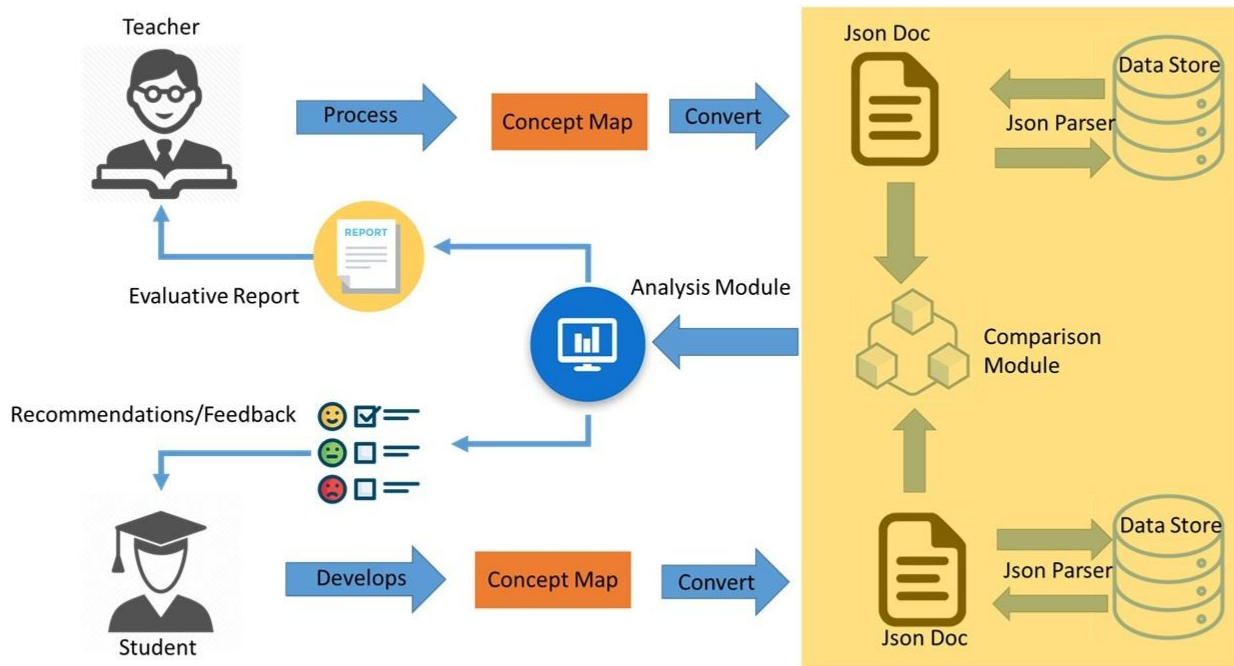


Figure 2 GSKCM system framework

The proposed framework facilitates automated evaluation of student-created CMs by comparing them with instructor-created master CMs, both represented in JSON format. A comparison module scores the conceptual alignment, while an analysis module interprets the results and presents them via a user interface. This enables instructors to assess understanding and students to reflect on their learning through structured feedback. The below points outline a structured CM evaluation framework wherein both instructor and student-generated CMs are converted into JSON format for automated comparison.

- The instructor creates a reference or master CM for a certain topic that is being studied.
- Students develop CM after studying the prescribed material/ topics discussed in the class.
- The master and reference CMs are both transformed into JavaScript Object Notation (JSON) documents.
- The comparison module of the tool extracts, compares, and scores the concepts embedded in the JSON documents. The proposed scoring algorithm has been implemented for scoring of concepts as well as calculating the overall score.
- The analysis module summarizes the findings from the user interface and evaluates the significance of the concepts that the students captured.
- The instructor assesses the evaluative report of the class or student using the user interface and renders a judgment on their understanding of the topic.
- The student evaluates the feedback and recommendations received for the CM they created and assesses their own understanding of the topic based on this feedback.

To create the reference CM, the instructor first generates the root concept on the CM canvas, as shown in *Figure 3*. Related concepts are created and placed in a basket or frame on the right side of the canvas. The instructor drags these concepts from the basket, linking them to the parent concept using the provided link labels, and saves the completed map to the database before scheduling the test. During the test, students see the root concept on the canvas and use the basket containing instructor-provided concepts. They drag and drop concepts to maintain structural relationships, select link labels, and rearrange them as needed. Once finalized, the student submits the CM, which is saved to the database.

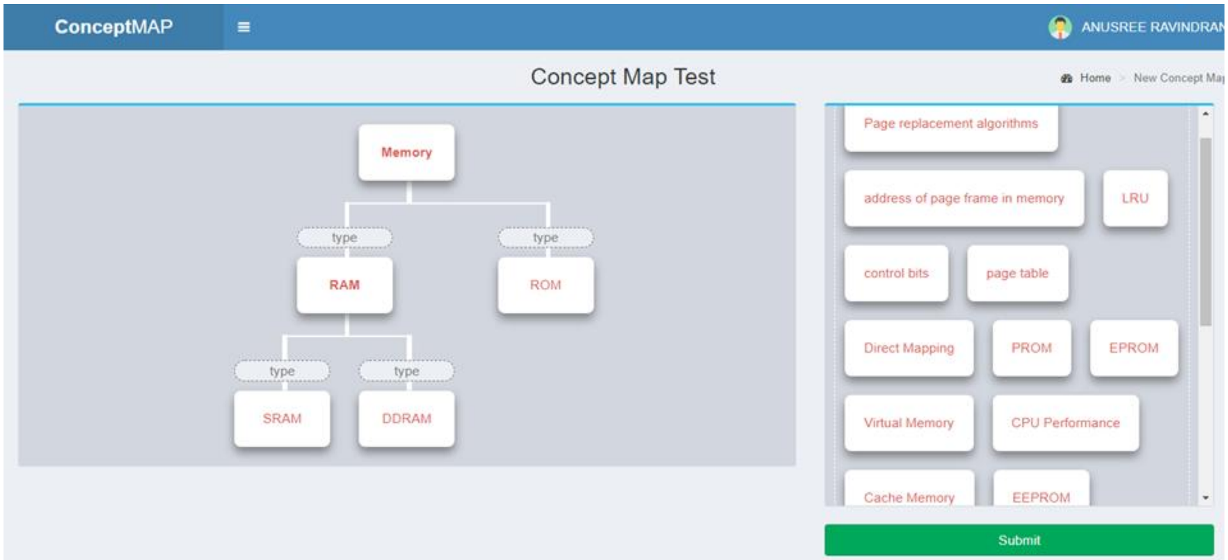


Figure 3 CM canvas

For comparing student CM and reference CM, the CM details are converted to JSON format while storing the CMs to the database. The JSON document for the master/reference CM is retrieved and compared with the JSON document of student CM. A JSON parser is used to extract the concepts and relations from the JSON document. The JSON document generated for the CM in *Figure 4* is shown in *Figure 5*. This JSON file contains all the required information on concepts and their associated relationships. The linking phrase between any two concepts represents the relation and comprehends these relations among different concepts, and is called "Concept Mapping". It also differentiates the concepts into different levels where the lower most level represents the depth of understanding the topic.

From *Figure 4*, it is observed that each and every concept is identified by four attributes:

- Data: refers to the concept name
- Link: refers to the linking phrase/label of the concept from the parent
- Parent: refers to the immediate parent of the concept
- Parent index: unique identifier for the parent concept

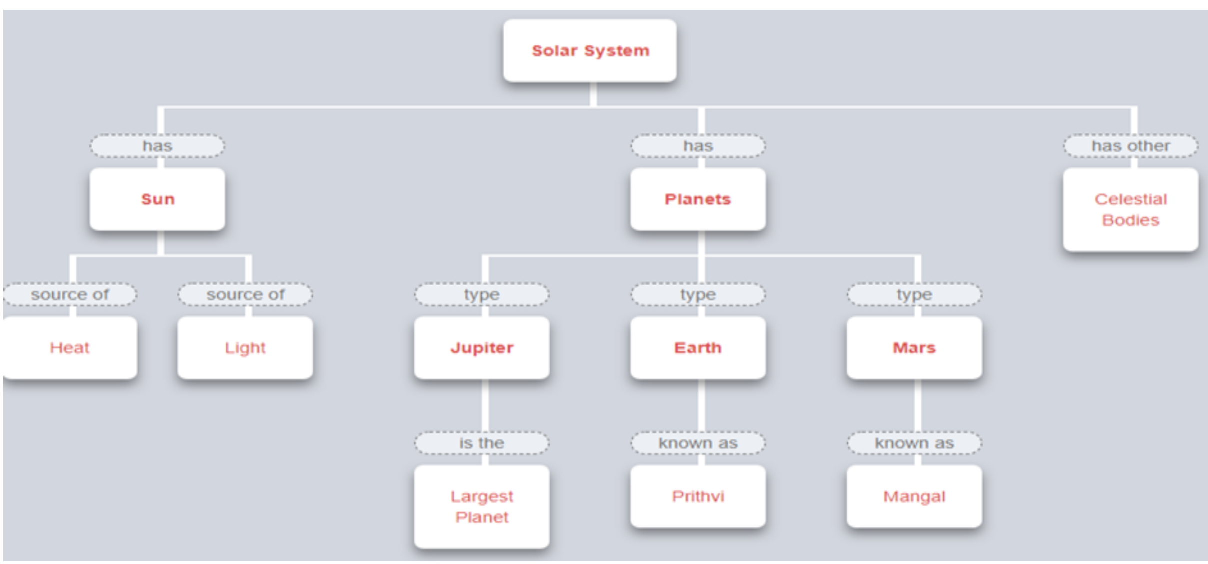


Figure 4 Sample CM

```
[{"data":"SolarSystem","link":"#","parent":"#","parentIndex":"#"},
{"data":"Sun","link":"has","parent":"Solar System","parentIndex":0},
{"data":"Heat","link":"source of","parent":"Sun","parentIndex":1},
{"data":"Light","link":"source of","parent":"Sun","parentIndex":1},
{"data":"Planets","link":"has","parent":"Solar System","parentIndex":0},
{"data":"Jupiter","link":"type","parent":"Planets","parentIndex":4},
{"data":"LargestPlanet","link":"is the","parent":"Jupiter","parentIndex":5},
{"data":"Earth","link":"type","parent":"Planets","parentIndex":4},
{"data":"Prithvi","link":"known as","parent":"Earth","parentIndex":7},
{"data":"Mars","link":"type","parent":"Planets","parentIndex":4},
{"data":"Mangal","link":"known as","parent":"Mars","parentIndex":9},
{"data":"CelestialBodies","link":"hasother","parent":"Solar System","parentIndex":0}]
```


Figure 5 JSON document of the CM

3.1CM scoring algorithm

The JSON documents of the master CM and the student CM are compared to grade the student's CM. The method incorporates the hierarchical structure of the concepts represented in the CM. This implies that the score assigned to each concept is determined on its proper placement within the parent hierarchy and its descendants. The student is believed to comprehend the concept if it is correctly placed in the hierarchy with regard to its parents and its descendants. The concept receives a full score in these circumstances. The method scores concepts based on the number of valid child concepts, the connection label, and the validity of all of the parent concepts in the hierarchy. If the parent hierarchical structure deviates from that of the master CM, the concept's score is proportionally reduced. The target/observed concept can be an internal node or a leaf node in the CM (*Figure 6*).

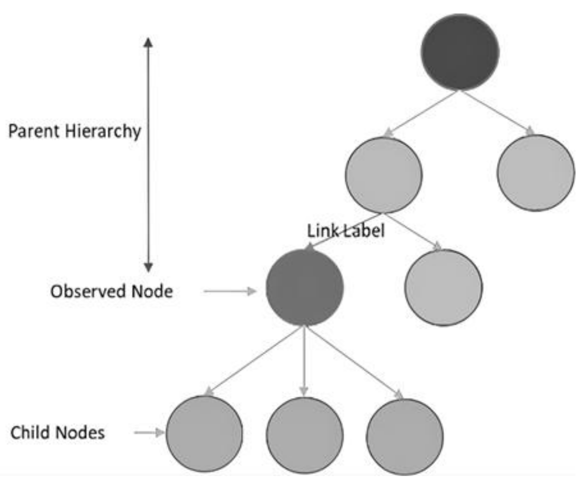


Figure 6 Types of nodes in a CM

The score of an internal node/concept is partitioned into:

◦ Incoming link label for the observed Concept **(Score Partition I)**
◦ Parent hierarchy of the observed concept **(Score Partition II)**
◦ Number of valid child nodes for the observed concept **(Score Partition III)**

The score of a leaf node/concept is partitioned into

◦ Incoming link label for the observed Concept **(Score Partition I)**
◦ Parent hierarchy of the observed concept **(Score Partition II & Score Partition III combined)**

Figure 7 shows an example of a Master CM on the topic Solar System with different types of nodes.

Scoring Partition of the concept" Mars" in the *Figure 7* is as follows:

- Sp1 refers to the correctness of the label on the link to the concept "Mars" (e.g., "types").
- Sp2 refers to the correctness of the parent hierarchy of "Mars" (e.g., "Planets" → "Solar System").
- Sp3 refers to the correctness of the child nodes of "Mars" (e.g., "Mangal").

Score partition I, illustrated in *Figure 8(a),* is based on the validity of the incoming link label for the observed concept, as defined in Equation 1.

$$sp1(c_i) = w_{sp1} \quad (1)$$

where, w_{sp1} is the weight assigned for the incoming label for the concept.

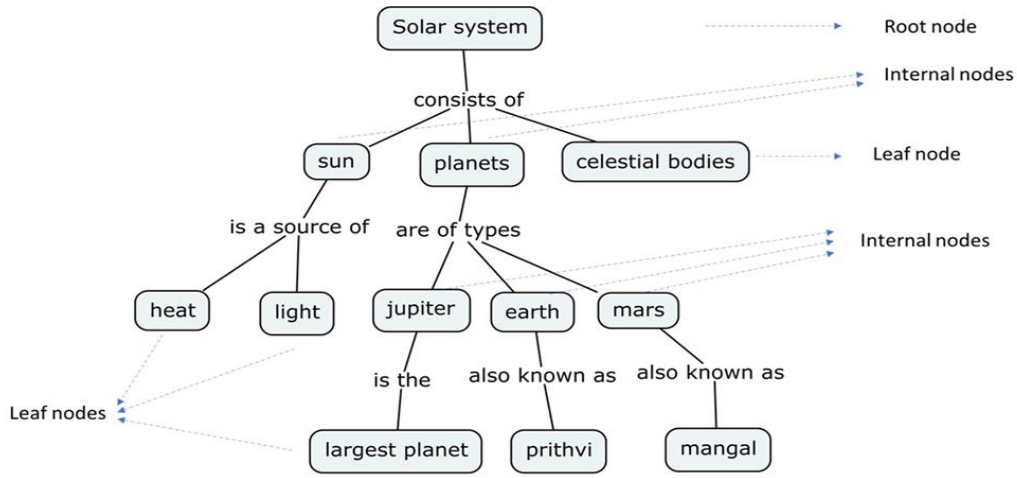


Figure 7 Master CM with different types of nodes

Score partition II as shown in *Figure 8(c)*, is based on the parent hierarchy of the observed concept in the master CM and is defined as follows:

$sp2(c_i) = \sum_{j=1}^{k}(MaxScore(p_j(c_i)) \times distdev(p_j(c_i))) \times count_{validParents}(c_i)/count_{totalParents}(c_i)$ (2)

where,

MaxScore($p_j(c_i)$) computes the proportion of the maximum score, a parent p_j of the observed concept c_i can attain based on its position in the parent hierarchy. Each parent node in the concept hierarchy is assigned a weight using an arithmetic progression, with the highest weight given to the immediate parent.

$MaxScore(p_j(c_i)) = w_{sp2} \times wp_{j\_mast}(c_i) / S$ (3)

where,

w_{sp2} is the weight assigned to the parent p_j of the parent hierarchy of the observed concept

w_{pj_mast}(ci) is initially set to k (the number of parent nodes) and is decremented for each subsequent parent node. The normalization factor S is calculated as 1 + 2 + 3 + ... + k.

Figure 8(b) illustrates the maximum score *MaxScore(p_j (c_i)* that each parent can achieve within the hierarchy assuming that score partition II contributes 60% to the total score of a concept.

To assess structural accuracy of the student CM, the upper distance—the distance between the observed parent concept and the root—and the lower distance—the distance between the observed parent and the observed concept—are calculated for both the student CM and the master CM, as depicted in *Figure 8(c)*. Based on their differences in distances, the distance deviation $distdev\left(p_j(c_i)\right)$, of the observed concept, c_i w.r.t to its observed parent p_j is computed.

$distdev(p_j(c_i)) = df/(distdif_{upper}(p_j(c_i) + distdif_{lower}(p_j(c_i))$ (4)

where,

df is the distance factor - a constant to adjust the variability of the deviation values

$distdif_{upper}$ $(p_j(c_i)$ is the upper distance difference between the observed parent p_j of c_i and the root concept

$distdif_{lower}$ $(p_j(c_i))$ is the lower distance difference between the observed parent p_j and the observed concept c_i

Thus, distdev(p_j (c_i)) computes the distance deviation factor of the parent p_j of the observed concept c_i of the student CM with respect to the master CM, as defined in Equation 4.

The average score of all parent nodes of the concept along with a penalty for incorrect parents $count_{validParents}(c_i)/count_{totalParents}(c_i)$, in the hierarchy are computed to get score partition II of the observed concept (C_i) as shown in Equation 2. The penalty is defined as the ratio of valid parent count of the observed concept($count_{validParents}(c_i)$ to the total parent count of the observed concept $count_{totalParents}(c_i)$.

Figure 8(d) depicts Score Partition III which is computed as follows:

$sp3(c_i) = w_{sp3} \times (count_{corr}ch(c_i)/(count\_ch(c_{i_{mast}}) + count_{incorr}ch(ci)))$ (5)

where,

w_{sp3} is the weight assigned for the correctness of the child nodes of the observed node.

$count_{corr}ch(c_i)$ is the number of valid child nodes under the observed concept in the student CM

$count\_ch(c_{i\text{-}mast})$ is the total number of valid child nodes given in the master CM for the observed concept

$count_{incorr}ch(c_i)$ is the number of invalid/missing child nodes, for the respective node as shown in Equation 5. The score for a concept ci is the sum of all score partitions as shown in Equations 6.

$$c_i\_score = sp1(c_i) + sp2(c_i) + sp3(c_i) \quad (6)$$

The total score for a student CM is $totalScore = \sum_{i=1}^{n}(ci\_score)/\ Count_{masternodes}$ (7)

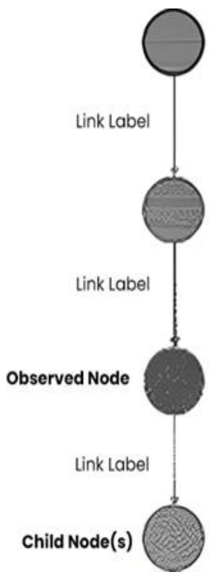


Figure 8(a) Score partition I

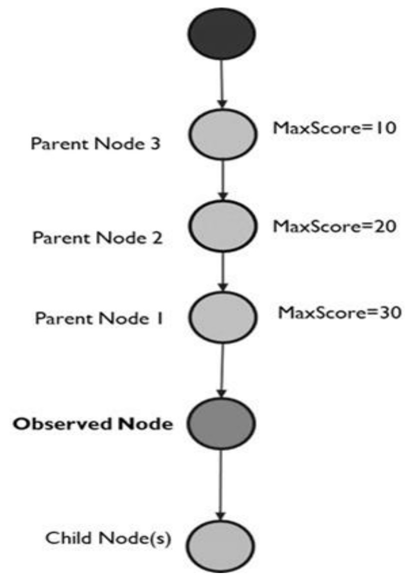


Figure 8(b) Score partition II-Maximum score for each parent concept in the master concept map

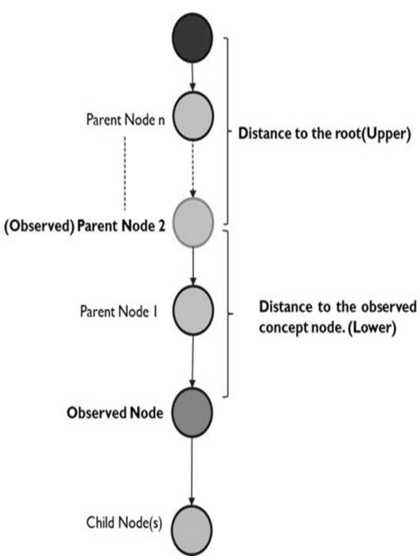


Figure 8(c) Score Partition II computation

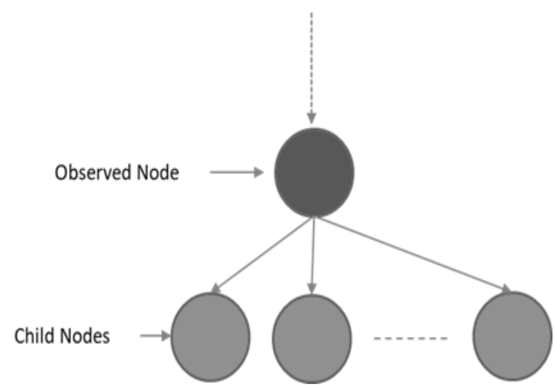


Figure 8(d) Score partition III

The scoring algorithm is shown below.

Scoring Algorithm

· ***Input:*** *masterTree =Master CM, tree=Student CM,* $\boldsymbol{w_{sp1}}, \boldsymbol{w_{sp2}}, \boldsymbol{w_{sp3}}$ refers to the weights assigned for incoming link label, parent concept hierarchy and child concepts. ***Output:*** Score of each concept c_i and the total score of student CM with reference to master CM.

Initialize totalScore and assign weights for each score partition wsp1, wsp2 & wsp3
For each node(c_i) in tree.
begin
  $count_{validParents}$ (c_i)= 0
  Find the masterNode in masterTree that matches the current node(c_i)
  if link labels of c_i and masterNode matches
  begin
    sp1(c_i)= w_{sp1} *// weight assigned for the incoming label for the concept*
  end
  If c_i is leaf node
  begin
    w_{sp2}= w_{sp2}+w_{sp3}
    w_{sp3}=0
  end
  $wp_{j\_mast}(c_i)=k$ *//k- no. of Parent nodes of c_i in the master CM*
  sp2(c_i)=0
    //computing the score of parent hierarchy w.r.t Maxscores of the parents with deviation in the student CM
  For each parentMasterNode P_j of masterNode
  begin
    Compute MaxScore($p_j(c_i)$) *// Maximum Score of parent p_j of c_i*
    $wp_{j\_mast}(c_i)$= $wp_{j\_mast}(c_i)$ -1
    For each parentNode of c_i
    begin
      if parentMasterNode data matches parentNode data)
      begin
        $count_{validParents}$**(c_i)**++
        compute distdev($p_j(c_i)$) //deviation in distance
        sp2(c_i)= **sp2(c_i)**+ $MaxScore\left(p_j(c_i)\right) \times$ distdev($p_j(c_i)$)
      end
    end
  end
    // average scores of parent hierarchy of the concept along with a penalty for incorrect parents in the hierarchy
    sp2(c_i) = **sp2(c_i)**)) × $count_{validParents}(c_i)/$ $count_{totalParents}(c_i)$
    Compute **sp3(c_i)**
    c_{i}_score = **sp1(c_i)+ sp2(c_i)+ sp3(c_i)** *//concept score of each concept*
    totalScore += c_{i}_score
end
totalScore= totalScore/$Count_{master}$ nodes *//total score of student CM*
Return totalScore

The algorithm calculates the total score of a student's CM by evaluating each concept node against a master CM. It checks for link matches (Score Partition I), evaluates parent hierarchy alignment (Score Partition II), and considers valid child node matches for internal nodes (Score Partition III). Leaf nodes skip Score Partition III, and weight adjustments are applied accordingly. The individual scores from all partitions are summed to compute the total score of the CM. The algorithm's complexity is assessed as O(n·d), where *n* represents the number of concepts and *d* denotes the maximum depth of the hierarchical tree formed by these concepts. Mathematically, the maximum possible height of a tree is $n - 1$, where n is the number of nodes (or concepts, in this case). In the context of hierarchical CMs, a singly linked tree implies that each concept has at most one child, forming a linear hierarchy. In such a scenario, the height of the tree is equivalent to the number of levels or concepts, resulting in a worst-case complexity of $O(n^2)$. However, this configuration is rarely observed in practical applications of hierarchical CMs. Real-world CMs typically exhibit branching and multiple interconnections, reflecting the complex relationships among concepts. As a result, CMs often have more balanced or branching hierarchies, which reduce the effective depth *d* and contribute to a more efficient computational structure.

3.2Validation of the scoring algorithm

Figure 9 shows the student generated CM on the topic Solar System for the master CM in *Figure 7*.

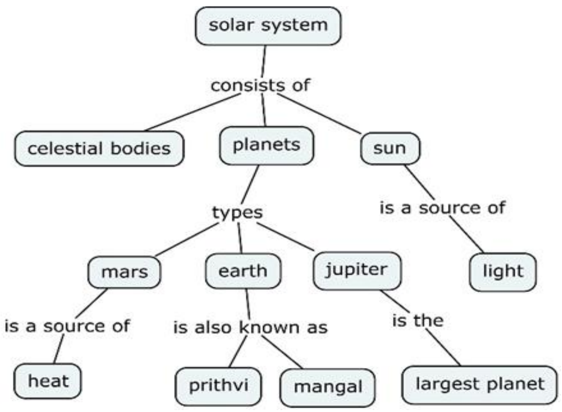


Figure 9 Student concept map

The *Table 1* describes how scores are computed for few selected concepts depicted in the student CM with reference to the master CM based on the scoring algorithm with 60% weightage of the score for parent hierarchy, 30% for validity of child nodes and 10% for the validity of incoming link label. Two cases can be considered here:

Case 1: Observed node is an internal node (e.g.: -concept "Mars" in *Figure 9*)

Here, the maximum score of each parent node of the observed node is computed from the master CM. Based on the deviation of the observed parent node in the student CM from the master CM, observed parent node score is computed. Averaging the score of parent nodes in the hierarchy of the observed concept contributes to sp2. Score with respect to the valid child nodes of the observed concept contributes to sp3 score. Validity of the link label to the observed concept contributes to sp1.

Case2: Observed node is a leaf node (e.g.: -concept "Mangal" in the *Figure 9*)

If the observed node is a leaf node, then only the score based on the parent hierarchy and incoming link label contributes to the concept score. The *Table 1* details the computation of the concept scores for both the cases specified. # indicates that the concept's parent is the root node. For leaf nodes, scores with respect to child nodes are marked as not applicable (NA).

Table 1 Concept Scores computed for few concepts of student CM based on the scoring algorithm

Concepts	**Master CM**				**Student CM**			
	Parent	**MaxScore($p_j(c_i)$)**	**Max Score w.r.t to child nodes (w_{sp3})**	**link score (w_{sp1})**	**Parent concept score attained sp2(c_i)**	**Score w.r.t to validity of child nodes sp3(c_i)**	**link score sp1(c_i)**	**Total**
Earth	Planet	60	30	10	60	15	10	**85**
heat (leaf node)	Sun	90	NA	10	0	NA	10	**10**
Mangal (leaf node)	Mars	60	NA	10	0	NA	10	**25**
	Planet	30	NA		15	NA		
Celestial Bodies (leaf node)	#	90	NA	10	90	NA	0	**90**
Planets	#	60	30	10	60	25	10	**95**
Mars	Planet	60	30	10	60	0	10	**70**

#-root concept, NA-Not Applicable

The comparison of CMs was conducted using different variations of student CMs to evaluate the accuracy of the scoring algorithm. The implemented algorithm assigns a score to each concept based on its correct placement within the hierarchical structure. If a concept is accurately positioned with respect to both its parent and child nodes, it is considered well understood by the student and receives a full score. Otherwise, the score is proportionally reduced based on the concept's deviation from its correct position in the parent hierarchy and the number of valid child nodes.

4.Results

An experiment utilizing the GSKCM framework was conducted with Class VIII students to evaluate the effectiveness of this innovative approach. A total of 35 students, following the NCERT syllabus, participated in the study. Prior to the assessment, students were instructed to independently study Chapter 10, titled "Sound", from their Science textbook (ISBN: 978-81-7450-812-6, 2022–23 edition). *Figure 10* presents the Master CMS created by the instructor for the chapter. This CM illustrates the interrelationships among 20 key concepts, with "Sound" serving as the root concept.

4.1Analysis of the CM "Sound" using GSKCM

The graph generated by the tool, as shown in *Figure 11*, illustrates the overall class average score of 61.53 for the topic "Sound." *Figure 12* presents the standard deviation of the test scores, offering insights into the distribution and variability of student performance across the class. In addition to the overall performance, the tool also displays the average scores for each individual concept within the topic.

From the graph, it is evident that certain concepts—such as loudness (34.72), thunder (39.69), and pitch (48.28)—have class average scores below 50%. These lower-performing areas highlight specific conceptual gaps among the students. Consequently, these concepts should be given additional emphasis and reinforcement during classroom reviews or instructional interventions to enhance overall understanding and retention.

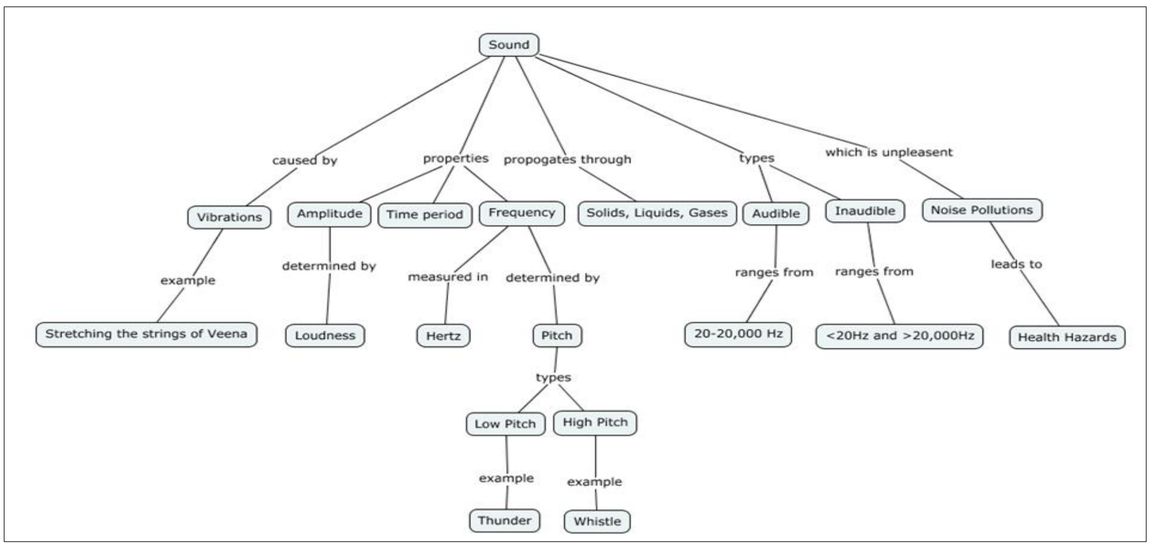


Figure 10 Master concept map on sound

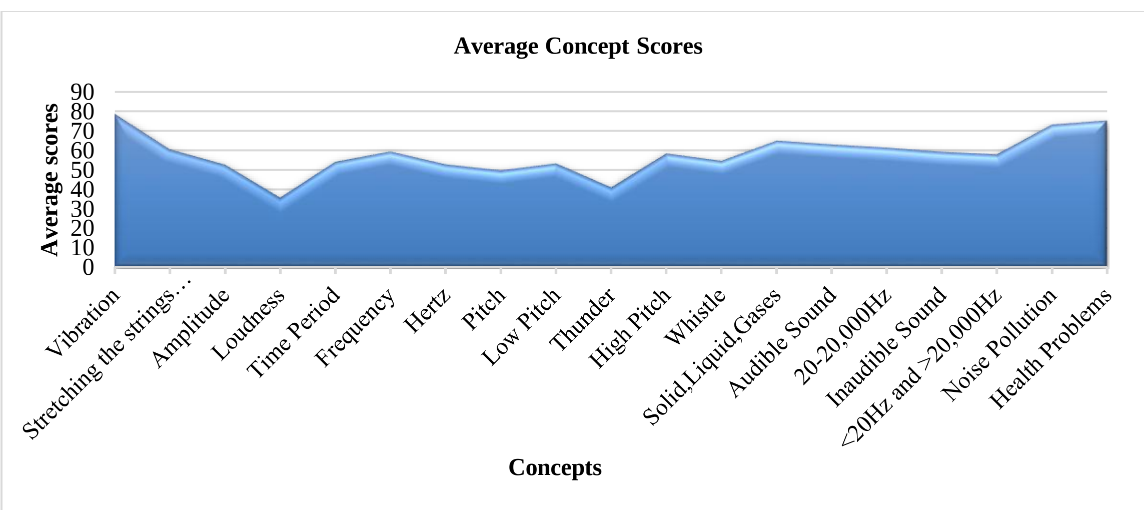


Figure 11 Class average of the test

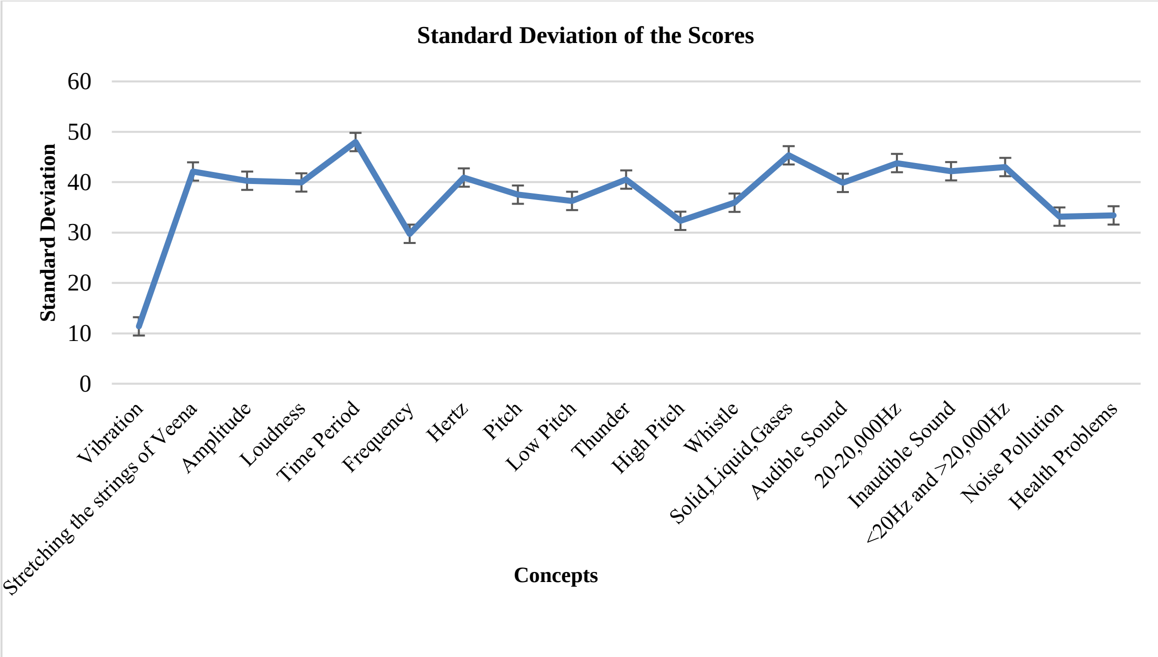


Figure 12 Standard deviation of the test scores

An analysis of the graph reveals that students demonstrated strong understanding in several key concepts, such as: "Sound is caused by vibrations, for example, stretching the strings of a Veena," "Sound propagates through solids, liquids, and gases," "Types of sound include audible and inaudible sounds, along with their respective frequency ranges," and "Unpleasant sounds contribute to noise pollution, which may lead to health issues."

However, certain concepts were identified as areas of weakness. These include: "Amplitude determines the loudness of a sound," "Frequency determines the pitch or shrillness of a sound," and "Different pitch types and their examples." While students recognized that the primary properties of sound are amplitude, frequency, and time period, they struggled to grasp the meaning and implications of each. For instance, many were unable to correctly associate examples with low and high pitch sounds. Similarly, they did not fully comprehend that amplitude influences loudness, frequency defines pitch or shrillness, and that frequency is measured in Hertz.

Given the overall class performance, it is evident that these foundational properties of sound need to be addressed in greater depth. Focused discussions and reinforcement activities targeting these specific weak concepts will support students in reconstructing and solidifying their understanding.

Additionally, individual student performance can be analyzed to identify specific misconceptions and gaps in knowledge. This diagnostic approach aids in personalized remediation, enabling students to rebuild their conceptual framework during the learning process [39]. *Figure 13* illustrates a comparison between the scores of the top three and bottom three performers in the class.

The framework also provides personalized feedback to each student, highlighting the specific concepts that require improvement based on their performance. *Figure 14* illustrates the concepts identified as needing further attention for a particular student.

4.2Clustering of students

To categorize students based on their level of comprehension of the topic, hierarchical clustering was employed. The Gap Statistic method indicated that the optimal number of clusters ranged between four and five. Based on this analysis, students were grouped into five distinct clusters. *Figure 15(a)* presents the cluster dendrogram, while *Figure 15(b)* displays the student groupings within each cluster. Additionally, *Figure 15(c)* illustrates the average concept scores for each of the five student clusters.

Figure 16 shows the standard deviation of each cluster.

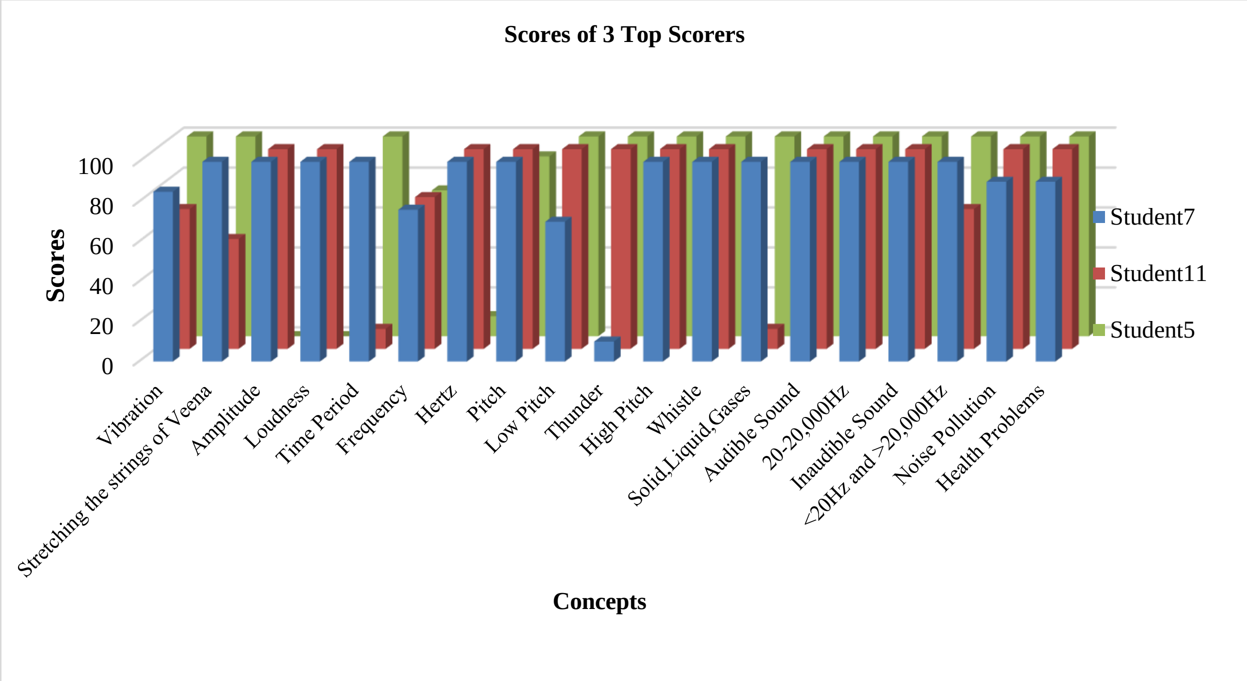


Figure 13 (a) Top scorers

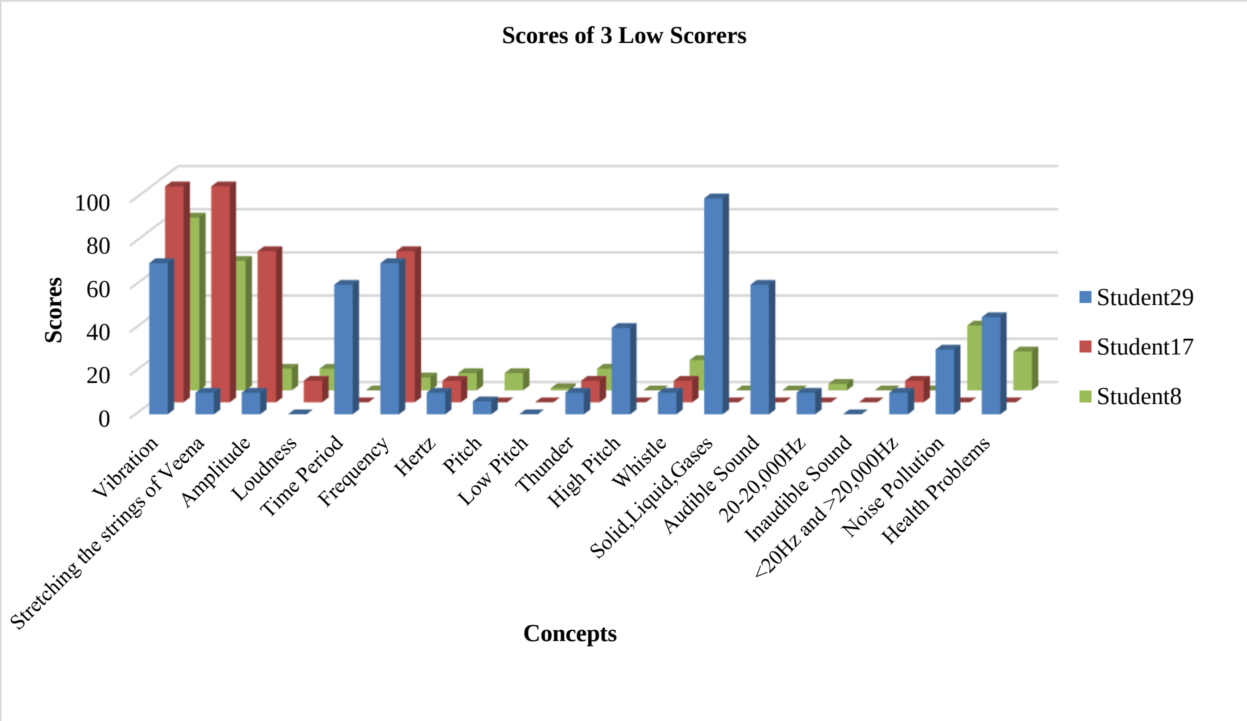


Figure 13 (b) Low scorers

Figure 14 Feedback interface

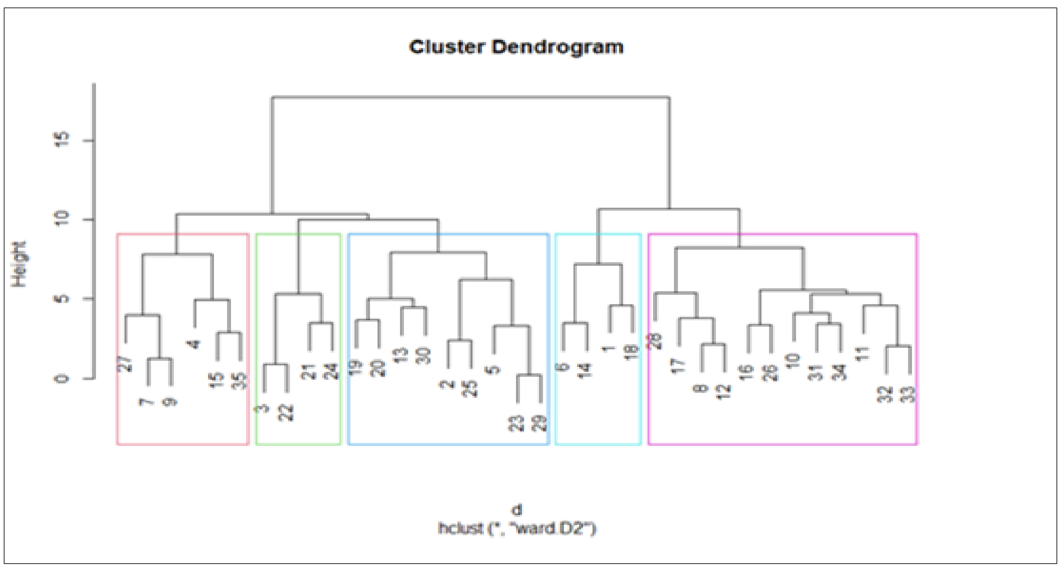


Figure 15(a) Cluster dendrogram

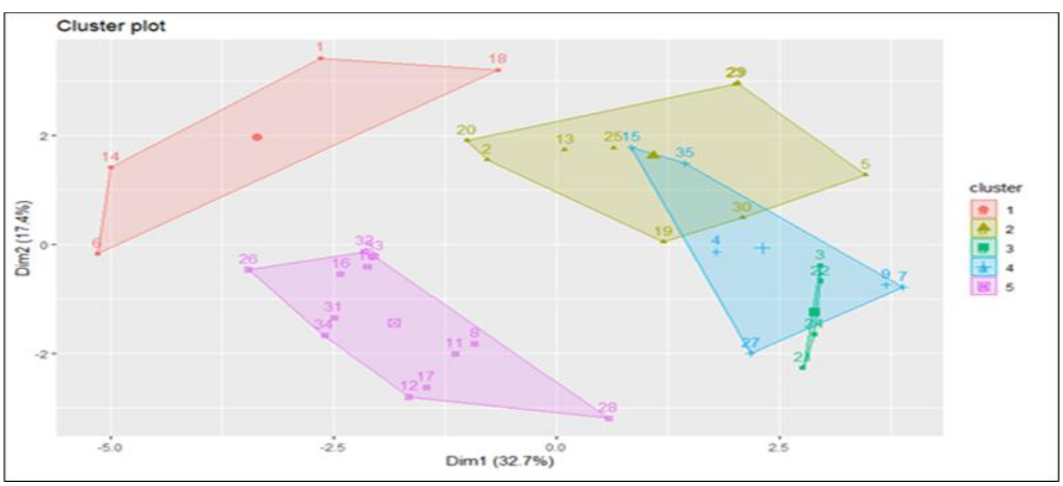


Figure 15(b) Cluster averages

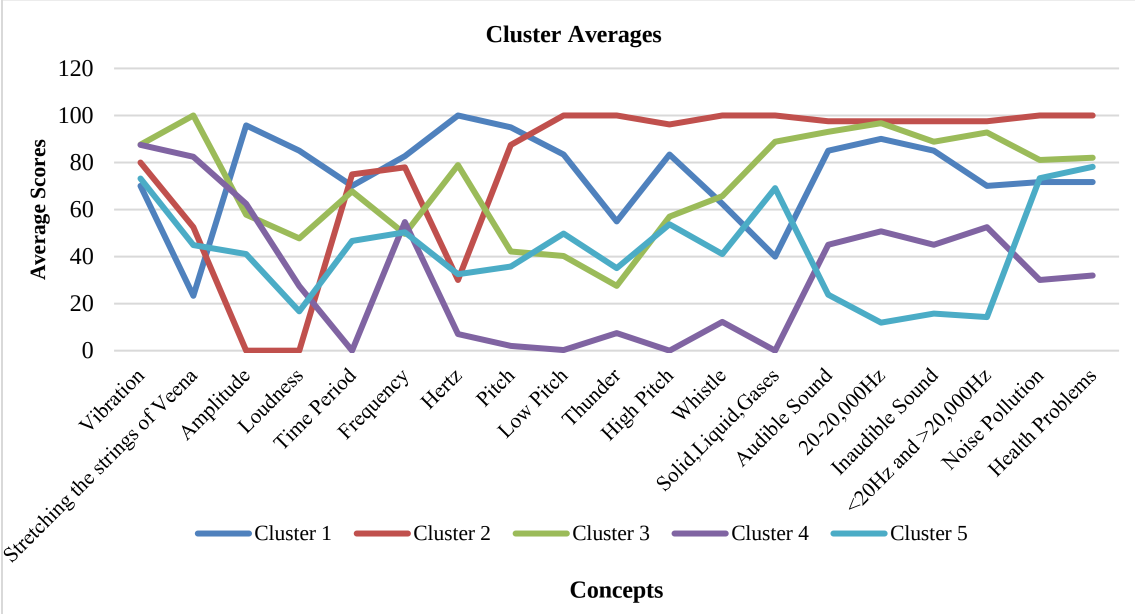


Figure 15(c) Cluster averages

Students in cluster1 possess a solid comprehension of the causes of sound and a fundamental understanding of sound properties, including Amplitude and Frequency. Their understanding of other concepts is significantly limited.

The students in cluster 2 have a low degree of grasp about the properties of sound, particularly in terms of their in-depth understanding of amplitude and frequency. They were able to comprehend the remaining ideas.

Cluster 3 students demonstrate a high degree of grasp in most topics, with the exception of amplitude, unit of frequency, and relating the example of stretching the strings of veena with the concept of vibration. Cluster 4 demonstrated a strong comprehension of all concepts except for the transmission medium of sound and failed to connect with the provided examples for the concept's vibration and low pitch. Cluster 5 exhibits a limited comprehension of most topics, with the exception of sound causes, sound propagation, and identification of noise hazards.

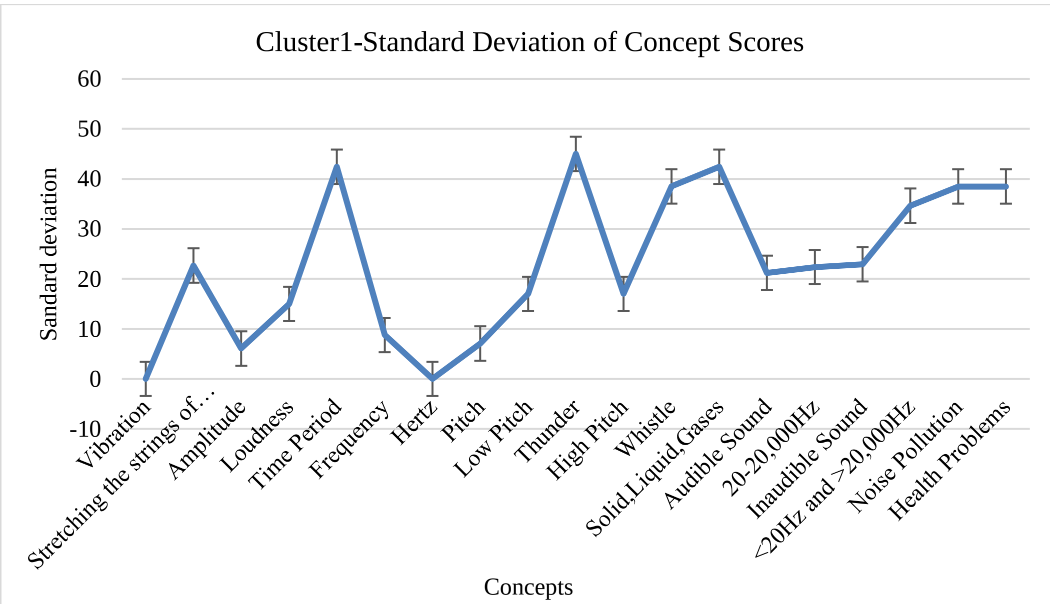
In summary, the level of comprehension for cluster 1 and 5 is low. Therefore, it is recommended to prioritize these students and provide them with additional attention to rectify their misconceptions and address any missing topics. Clusters 3 and 4 are exhibiting satisfactory performance, nevertheless, they require focused attention on the few concepts that were not fully grasped. Each cluster of students were given special remedial sessions and was asked to reiterate the construction of CMs with the practice module of the GSKCM tool. The *Figure 17(a),17(b)* presents the test scores of two groups of students in their formative assessment on the chapter "Sound." One group, the experimental group, practiced learning using CMs, while the other group, the control group, did not use CMs for learning. The *Figure 17* indicates that the experimental group outperformed the control group. improve language.

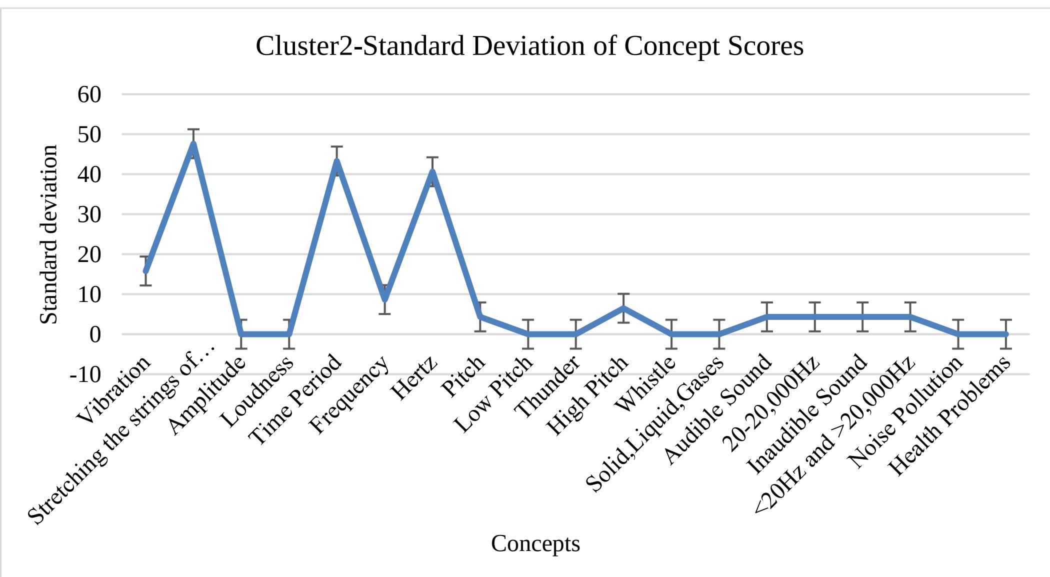
The educational implications of clustering students lie in the ability to tailor instruction to meet diverse learning needs. By identifying clusters, educators can group students with similar learning profiles, allowing for more personalized teaching strategies. This can help address specific challenges faced by each group, such as providing additional support to struggling students or offering advanced materials to high achievers. Clustering also enables teachers to monitor group progress and adjust their teaching methods accordingly, fostering a more inclusive and effective learning environment. Ultimately, understanding these clusters can lead to more observed interventions, improving overall student engagement and achievement.

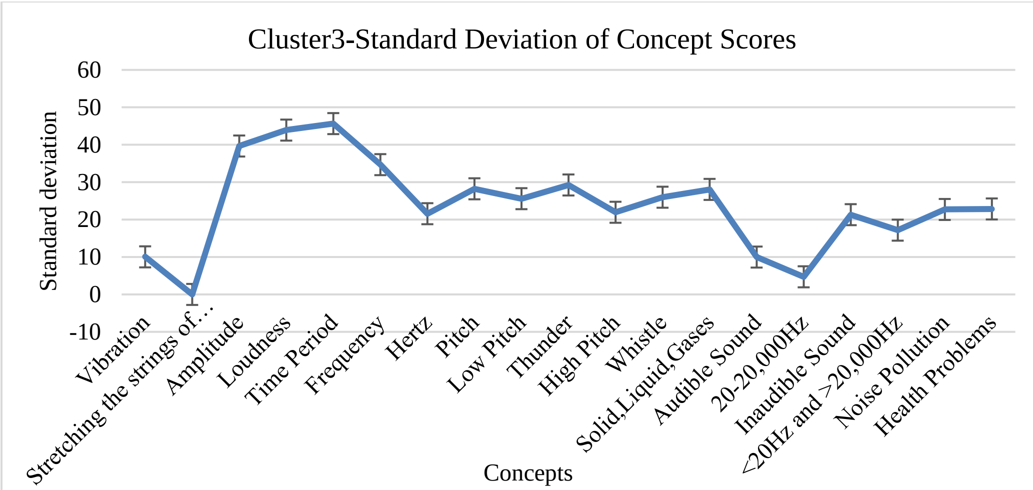
Several strategies can be employed to address these misconceptions for each cluster groups.

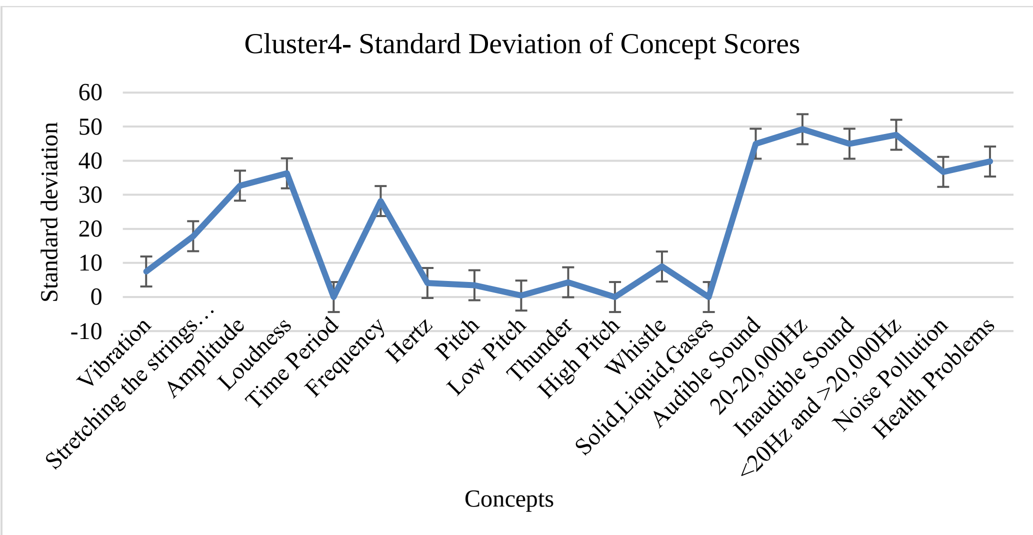
1. Encourage students to reflect on their CMs by asking them to explain their reasoning for the connections they made. This can help uncover and

address underlying misconceptions. Teachers can provide corrective feedback to guide students toward more accurate understanding [40].
2. Explicitly teach common misconceptions related to the subject matter before students create their CMs. Addressing these misconceptions upfront can help prevent them from being incorporated into the maps [41]. Design activities that specifically observed common misconceptions by asking students to create maps on those topics. Following this, the teacher can lead a discussion or provide resources to correct any errors, guiding students toward a more accurate understanding.
3. Use technology-based support tools that guide students through the process of map creation, offering hints or suggestions when a misconception is detected. Such scaffolding helps students to construct more accurate maps and internalize correct concepts.
4. Encourage students to revisit and revise their CMs over time, allowing them to correct misconceptions as they deepen their understanding. This iterative approach reinforces learning and helps students refine their conceptual frameworks [42].









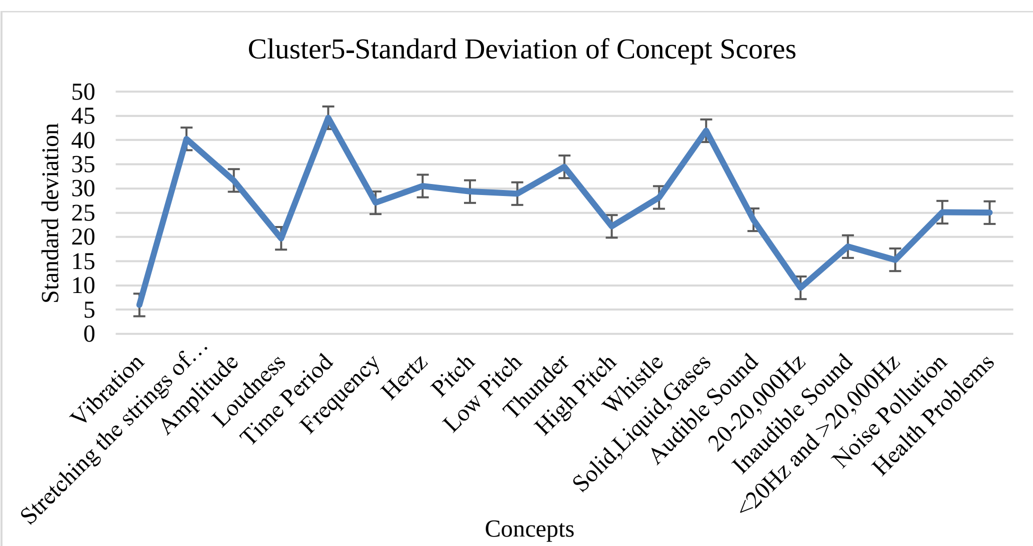


Figure 16 Standard deviation of concept scores in each cluster

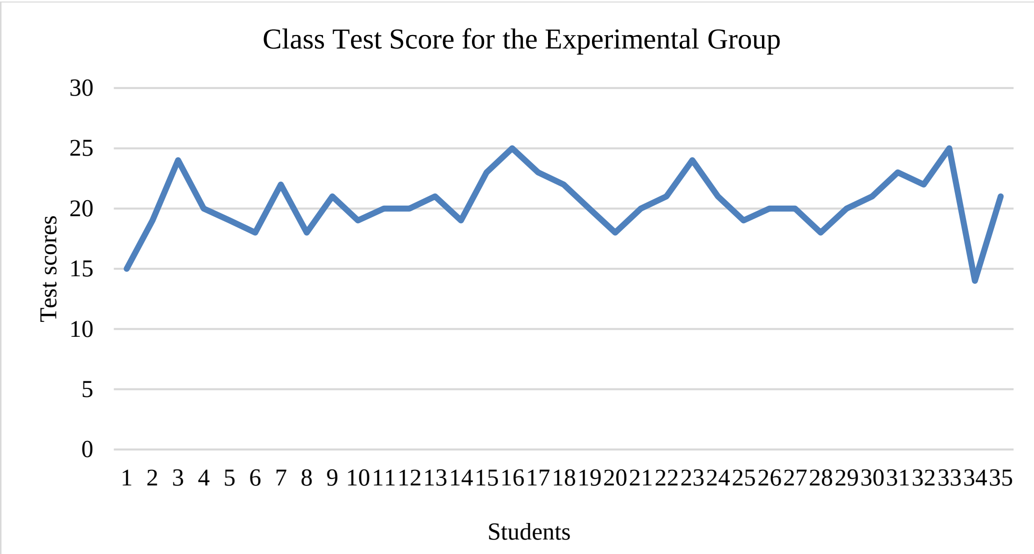


Figure 17(a) Test score of experimental group

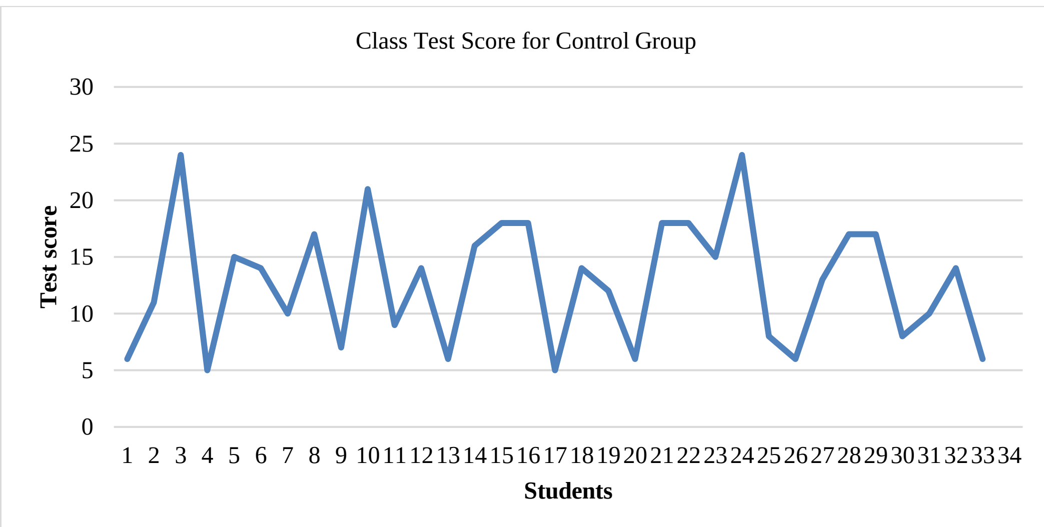


Figure 17(b) Test score of control group

4.3Mapping the students score to the LOs

To evaluate the overall performance of the class, student scores were aligned with the intended LOs. *Table 2* presents the mapping of five LOs for the chapter "Sound," as defined by the instructor.

Table 2 Mapping LOs to related concepts in the master CM

LOs	Related Concepts in the CM	Average Score/ LO
Understand the causes of Sound	Vibration, Stretching the strings of Veena	68.42
Analyse the properties of Sound	Amplitude, Time period, Frequency, Pitch, High Pitch, Low Pitch, Thunder, Whistle, Hertz	54.67
Classify Sound types	Audible Sound, Inaudible Sound,20-20,000Hz, <20 and >20,000Hz	62.9
Recognize noise hazards	Noise Pollution, Health problems	80.25
Understand Sound Propagation	Solid-Liquid-Gases	70.93

Table 2 clearly indicates that the majority of students demonstrate weakness in understanding Sound Properties, highlighting the need for a more in-depth exploration of this concept.

It is important to note that a student's score in a conventional written test should not be directly compared with their CM score. This is because written test scores often reflect a student's ability to memorize subject content, whereas CM scores are indicative of the student's structural understanding and conceptual knowledge alone.

5.Discussion

The GSKCM framework effectively automates the evaluation of CMs by validating the structural hierarchy and conceptual relationships, providing accurate assessments of students' understanding. The algorithm successfully identifies weaker concepts, unknown concepts, and misconceptions, and clusters students based on their scores for observed remediation. The results indicate that this method is both efficient and reliable in assessing student knowledge.

The findings suggest that the GSKCM framework enhances the precision and consistency of CM evaluations by focusing on structural validation, which traditional methods may overlook. The ability to cluster students based on their conceptual understanding allows for more personalized and effective remediation strategies, supporting diverse learning needs. The implementation of GSKCM in educational settings can significantly improve the assessment process, particularly in large classes where manual evaluation is impractical. By providing detailed, structure-based feedback, GSKCM not only facilitates a better understanding of student learning but also enables educators to tailor their teaching strategies more effectively. This approach could lead to improved educational outcomes by addressing individual learning gaps more systematically. Overall, GSKCM offers a robust solution for CM assessment, combining automation with a focus on structural integrity, which is essential for a comprehensive understanding of student knowledge.
The GSKCM framework is compared with several established scoring algorithms and frameworks, including the Relational, Holistic, Waterloo, and Topological scoring methods. *Table 3* presents a comparative analysis of these approaches. The GSKCM scoring method is automated, enabling the evaluation of both individual concepts and holistic CMs. However, it does not support the creation and evaluation of CMs from scratch. Instead, students are provided with a predefined list of concepts related to the topic for CM construction. In contrast, the Relational, Holistic, Waterloo, and Topological scoring methods rely on manual evaluation. These methods involve manual validation of concepts, as well as validation of both hierarchical and non-hierarchical structures within the CMs. Since the entire process is conducted manually, these methods support both the creation and holistic evaluation of CMs. *Table 4* compares the GSKCM framework with similar frameworks. The frameworks compared are fuzzy based framework, artificial intelligence-based student learning evaluation (AISLE) framework, Markov Chain framework, ML based automatic evaluation framework and graph neural network (GNN) framework.

The GSKCM framework distinguishes itself through its high level of automation and targeted educational utility. It supports automated validation of concepts and hierarchical taxonomy, offers both individual concept scoring and holistic CM evaluation, and includes visualization of results. A key strength is its ability to segment students based on their concept scores, allowing for focused remediation. However, GSKCM has certain limitations—it does not support the evaluation of non-hierarchical CMs and restricts students to constructing CMs using a predefined list of concepts, thus limiting creativity and free-form mapping.

In contrast, other frameworks such as AISLE and the Markov Chain approach support the creation and evaluation of CMs from scratch and can handle non-hierarchical structures. These frameworks rely on manual validation and often evaluate understanding across multiple structural levels (gist, support, detail). While the Fuzzy-based and ML-based frameworks incorporate automation, they are generally more limited in scope—focusing on proposition-level validation or lacking detailed structural assessments. GNN Frame work automatically learn rich structural features from concept maps without manual feature engineering, but disadvantage is they do not evaluate concept correctness, hierarchy, or provide detailed semantic feedback. Compared to these, GSKCM offers a more comprehensive and student-centered approach, particularly effective for formative assessment and targeted instructional support.

The current framework ensures a complete tool for hierarchical CM construction with a given list of concepts and evaluating it by validating the hierarchical taxonomy along with the concept and link label (proposition)validation facilitating concept scoring at the individual level as well as at the holistic level which is first of its kind.

Table 3 Comparison of the GSKCM scoring algorithm with other scoring algorithms

Characteristics	GSKCM scoring algorithm	Relational/Holistic/Waterloo/ topological scoring method[9, 30, 31]
Concept validation	Automated	Manual
Validating the hierarchical taxonomy	Automated	Manual
Individual concept scoring	Yes	No
Holistic CM scoring	Yes	Yes
Able to evaluate CMs created from scratch	No	Yes
Evaluating Non-hierarchical CM	No	Yes

Table 4 Comparison of the current framework with other related frameworks

Characteristics	GSKCM framework	Fuzzy based framework [16, 32]	AISLE framework[14]	Markov chain frame work[15]	ML-based automatic evaluation[43]	GNN framework[44]
Concept validation	Automated	Automated	Manual	Manual	Proposition validation automated rather than concept	Concept Meaning are not Validated
Validating the hierarchical taxonomy	Automated	Automated	Validated at three levels. Further levels are iterated recursively to either level 2 or 3	Validated at three levels-gist, support and detail level	No	No explicit hierarchy checking
Individual concept scoring	Yes	No	Random scores are based on the three levels.	No	Proposition is scored	No
Holistic CM scoring	Yes	Yes	Probability distribution-based scoring system evaluates CM quality. The emphasis is placed on an individual's comparative comprehension of the topic in relation to other students.	Length of the bar chart at gist, support and detail levels determine the quality of the CM	No	Partial-No traditional "grading" but grouping for educators to assess understanding levels.
Able to evaluate CMs created from scratch	Students are given a list of concepts for mapping	No	Yes	Yes	No	Yes
Evaluating non-hierarchical CM	No	Yes	No	Yes	No	Yes
Visualisation of Results	Yes	Yes	Yes	Yes		Only cluster visualizations
Segmentation of students based on their concept scores for remediation	Yes	No	No	No	No, suggested as future step	No - only CMs are clustered

Table 5 Comparison of the current framework with other popular educational frameworks

Aspect	**GSKCM framework**	**Constructivist learning environments (CLEs)**[45]	**Flipped classroom model**[46]	**Bloom's taxonomy-based frameworks**[47]
Student Engagement	Student engagement through Concept Mapping and immediate feedback, helping students stay actively involved.	Promotes active, student-centered learning but may not cater to the needs of all types of students, leading to varying levels of engagement	High engagement through interactive class activities, but may not offer the continuous feedback loop found in GSKCM	Engages students through activities observing different cognitive levels, but may lack the visual, immediate feedback mechanisms of GSKCM
Learning	Enhances deep learning by encouraging iterative refinement of CMs, aiding in the correction of misconceptions and retention	Effective in fostering critical thinking and problem-solving, though LOs may be less consistent without structured guidance	Improves understanding and retention by allowing self-paced learning and application, but lacks the continuous guided feedback central to GSKCM	Promotes a wide range of cognitive skills, from basic recall to higher-order thinking, but may not provide the clear visual outcomes that GSKCM offers

Table 5 compares the GSKCM framework with other widely-used educational frameworks, emphasizing its distinct advantages in providing structured guidance and iterative feedback. It suggests that integrating existing frameworks with GSKCM could lead to significantly improved results. Such integration would combine the strengths of both approaches, enhancing student engagement and LOs by offering a more comprehensive and supportive educational experience. This alignment could cater to a wider range of student needs, fostering deeper understanding and more effective knowledge retention.

This framework aligns with several established educational theories such as constructivist learning theories, metacognitive practices, formative assessment practices etc. The GSKCM framework is closely tied to constructivist learning theories, which emphasize the importance of learners actively constructing their own understanding of concepts. GSKCM facilitates this process by enabling students to create and iteratively refine their CMs, thereby constructing and restructuring their knowledge in a meaningful way. GSKCM also supports metacognitive practices, encouraging students to reflect on their learning process, evaluate their understanding, and make necessary adjustments. Recent educational research highlights the importance of metacognition in enhancing LOs. By providing a clear framework for reflection and self-assessment, GSKCM aligns with these findings and can significantly improve students' ability to manage their own learning [48]. The GSKCM framework is well-suited for formative assessment practices, which focus on providing ongoing feedback to support student learning. GSKCM's guided scoring system provides detailed, iterative feedback that can help students refine their understanding continuously [28].

Constructive alignment emphasizes the need for coherence between learning objectives, teaching activities, and assessment tasks. GSKCM fits well within this framework by ensuring that the process of Concept Mapping and the associated scoring are directly aligned with the desired LOs. This alignment ensures that students are engaged in activities that directly contribute to their understanding and mastery of the subject matter [49]. The GSKCM framework should be refined to better accommodate diverse and cyclic CM structures. This will ensure that the algorithm can accurately assess creative or unconventional maps without compromising the validity of the evaluation. GSKCM could be integrated with other assessment methods, such as formative assessments or peer evaluations, to provide a more comprehensive view of student learning. This holistic approach would enrich the feedback provided to students and educators.

5.1Limitations

GSKCM is primarily designed for hierarchical CMs, which is a key limitation as the scoring algorithm assesses only those that follow a strict top-down structure. This limits its effectiveness in assessing non-hierarchical or cyclic CMs, which can better represent certain types of knowledge or learning domains (e.g., systems thinking). Moreover, in real-world contexts, CMs often exhibit non-hierarchical or cyclic structures, where concepts are interconnected in more complex ways, reflecting the true nature of knowledge. Another limitation is that the framework relies on a predefined list of concepts from the instructor, which may constrain students' creativity and limit their ability to explore novel connections or concepts. This restricts the evaluation of their deeper understanding and critical thinking abilities. The framework depends on a predefined list of concepts provided by the instructor, which may constrain students' creativity and limit their ability to explore novel connections or concepts not initially identified by the instructor.

A complete list of abbreviations is listed in *Appendix I*.

6.Conclusion and future work

The proposed framework is unique for constructing hierarchical CMs with a provided list of concepts, validating the hierarchical taxonomy, concept and link labels, and facilitating individual and holistic concept scoring. The CM construction and evaluation tool was tested in an authentic classroom setting. The performance report helped the teacher identify weaknesses in the collective and individual understanding of students. Students were clustered according to their comprehension of each concept. An analysis of the clusters identified strong and weak concepts within each group, enabling the provision of observed remediation. Concepts are aligned to the LOs of the chapter. Concept scores are then used to evaluate the LOs for the class and individual students based on their test performance. In conclusion, GSKCM framework assesses students' understanding and facilitates observed remediation based on strengths and weaknesses, improving educational teaching and learning.

A modified tool can be designed to help students create CMs from scratch, allowing them to generate concepts related to a topic and assign meanings using their own link labels. To resolve the ambiguity in the students' concept names and linking phrases, text mining and AI techniques may also be implemented.

Acknowledgment
None.

Conflicts of interest
The authors have no conflicts of interest to declare.

Data availability
The data considered in this study were collected from Rajagiri Public School, Kochi, Kerala, India involving students of Class VIII. The data are not publicly available. However, the data may be provided by the corresponding author upon reasonable request.

Author's contribution statement
Sunu Mary Abraham: Conceived, designed and implemented the system, conducted experiments, collected and analysed data, and drafted the manuscript. **G. Sudhamathy**: Supervision, guided the research direction, and reviewed the manuscript.

References

[1] Novak JD, Cañas AJ. The theory underlying concept maps and how to construct them. Florida Institute for Human and Machine Cognition. 2006; 1(1):1-31.

[2] Jackson A, Barrella E, Bodnar C. Application of concept maps as an assessment tool in engineering education: systematic literature review. Journal of Engineering Education. 2024; 113(4):752-66.

[3] Hartmeyer R, Stevenson MP, Bentsen P. A systematic review of concept mapping-based formative assessment processes in primary and secondary science education. Assessment in Education: Principles, Policy & Practice. 2018; 25(6):598-619.

[4] Rohles B, Backes S, Fischbach A, Amadieu F, Koenig V. Creating positive learning experiences with technology: a field study on the effects of user experience for digital concept mapping. Heliyon. 2022; 8(4);1-21.

[5] Cañas AJ, Novak JD. Concept mapping using CmapTools to enhance meaningful learning. In knowledge cartography: software tools and mapping techniques 2014 (pp. 23-45). London: Springer London.

[6] Zagaar M, Chen W. Assessment for deeper understanding using concept maps: lessons learned from flipped teaching of pharmacology. Medical Science Educator. 2022; 32(6):1289-97.

[7] Alt D, Naamati-schneider L. Health management students' self-regulation and digital concept mapping in online learning environments. BMC Medical Education. 2021; 21:1-5.

[8] Tan S, Erdimez O, Zimmerman R. Concept mapping as a tool to develop and measure students' understanding in science. Acta Didactica Napocensia. 2017; 10(2):109-22.

[9] Cetın B, Guler N, Sarıca R. Using generalizability theory to examine different concept map scoring methods. Eurasian Journal of Educational Research. 2016; 16(66):211-28.

[10] Chang CY, Yang JC. Concept mapping in computer-supported learning environments: a bibliometric analysis. Interactive Learning Environments. 2023; 31(10):6678-95.

[11] Kapuza A. How concept maps with and without a list of concepts differ: the case of statistics. Education Sciences. 2020; 10(4):1-13.

[12] Prasetya DD, Pinandito A, Hayashi Y, Hirashima T. Analysis of quality of knowledge structure and students' perceptions in extension concept mapping. Research and Practice in Technology Enhanced Learning. 2022; 17(1):1-25.

[13] Shao Z, Li Y, Wang X, Zhao X, Guo Y. Research on a new automatic generation algorithm of concept map based on text analysis and association rules mining. Journal of Ambient Intelligence and Humanized Computing. 2020; 11(2):539-51.

[14] Jain GP, Gurupur VP, Schroeder JL, Faulkenberry ED. Artificial intelligence-based student learning evaluation: a concept map-based approach for analyzing a student's understanding of a topic. IEEE Transactions on Learning Technologies. 2014; 7(3):267-79.

[15] Gurupur VP, Jain GP, Rudraraju R. Evaluating student learning using concept maps and markov chains. Expert Systems with Applications. 2015; 42(7):3306-14.

[16] Dias SB, Dolianiti FS, Hadjileontiadou SJ, Diniz JA, Hadjileontiadis LJ. FISCMAP: a fuzzy logic-based quality of concept mapping modelling approach fostering reflective feedback. In proceedings of the 7th international conference on software development and technologies for enhancing accessibility and fighting info-exclusion 2016 (pp. 293-300). ACM.

[17] Steiner CM, Albert D. Validating domain ontologies: a methodology exemplified for concept maps. Cogent Education. 2017; 4(1):1-39.

[18] Shou Z, Lu X, Wu Z, Yuan H, Zhang H, Lai J. On learning path planning algorithm based on collaborative analysis of learning behavior. IEEE Access. 2020; 8:119863-79.

[19] Wang B, Giabbanelli PJ. Identifying informative features to evaluate student knowledge as causal maps. International Journal of Artificial Intelligence in Education. 2024; 34(2):301-31.

[20] Li Y, Shao Z, Wang X, Zhao X, Guo Y. A concept map-based learning paths automatic generation algorithm for adaptive learning systems. IEEE Access. 2018; 7:245-55.

[21] Siew CS. Using network science to analyze concept maps of psychology undergraduates. Applied Cognitive Psychology. 2019; 33(4):662-8.

[22] Kroeze KA, Van DBSM, Veldkamp BP, De JT. Automated assessment of and feedback on concept maps during inquiry learning. IEEE Transactions on Learning Technologies. 2021; 14(4):460-73.

[23] Yang LP, Xin T, Luo F, Zhang S, Tian XT. Automated evaluation of the quality of ideas in compositions based on concept maps. Natural Language Engineering. 2022; 28(4):449-86.

[24] De RKE, Schaap H, Van LAM, Kral MM, Meijer PC. A literature review of open-ended concept maps as a research instrument to study knowledge and learning. Quality & Quantity. 2022: 1-35.

[25] Dmoshinskaia N, Gijlers H, De JT. Giving feedback on peers’ concept maps as a learning experience: does quality of reviewed concept maps matter? Learning Environments Research. 2022; 25(3):823-40.

[26] Su F, Zou D. A comparative review of technology-assisted and non-technology concept mapping-based language learning. International Journal of Educational Research Open. 2024; 6:1-10.

[27] Chatzi AV, Kourousis KI. Are concept map exam papers reliable as assessment tools in nursing education? a quantitative research pilot study. Teaching and Learning in Nursing. 2023; 18(2):293-8.

[28] Abd HH, Hamzah MI, Zulkifli H. A systematic review on the formative assessment practice in teaching and learning in secondary school. International Journal of Evaluation and Research in Education. 2024; 13(2):1173-83.

[29] Veiga F, Gil-del-val A, Iriondo E, Eslava U. Validation of the use of concept maps as an evaluation tool for the teaching and learning of mechanical and industrial engineering. International Journal of Technology and Design Education. 2025; 35(1):383-401.

[30] Bhatia S, Bhatia S, Ahmed I. Automated waterloo rubric for concept map grading. IEEE Access. 2021; 9:148590-8.

[31] Deshpande P, Ahmed I. Topological scoring of concept maps for cybersecurity education. In proceedings of the 50th technical symposium on computer science education 2019 (pp. 731-7). ACM.

[32] Al DM, Newbury P. Adaptive e-learning system using fuzzy logic and concept map. World Academy of Science, Engineering and Technology International Journal of Educational and Pedagogical Sciences. 2018; 21:1-8.

[33] Dahir MA, Qasim SA, Ahmed I. Cronus: an automated feedback tool for concept maps. IEEE Access. 2021; 9:119564-77.

[34] Ho VW, Harris PG, Kumar RK, Velan GM. Knowledge maps: a tool for online assessment with automated feedback. Medical Education Online. 2018; 23(1):1-9.

[35] Costin V. Online knowledge maps with automated feedback in education and assessment. Doctoral Dissertation, Costin, Veronica, Medical Sciences, Faculty of Medicine, UNSW. 2018.

[36] Mclinden D. Concept maps as network data: analysis of a concept map using the methods of social network analysis. Evaluation and Program Planning. 2013; 36(1):40-8.

[37] Goldman AW, Kane M. Concept mapping and network analysis: an analytic approach to measure ties among constructs. Evaluation and Program Planning. 2014; 47:9-17.

[38] Sdravopoulou K, Muñoz GJM, Hidalgo-ariza MD. Network analysis for learners’ concept maps while using mobile augmented reality gaming. Applied Sciences. 2021; 11(21):1-15.

[39] Santos VD, Souza ÉF, Felizardo KR, Vijaykumar NL. Analyzing the use of concept maps in computer science: a systematic mapping study. Informatics in Education. 2017; 16(2):257-88.

[40] Moon JA. Reflection in learning and professional development: theory and practice. Routledge; 2013.

[41] Smith IIIJP, Disessa AA, Roschelle J. Misconceptions reconceived: a constructivist analysis of knowledge in transition. The Journal of the Learning Sciences. 1994; 3(2):115-63.

[42] Eshuis EH, Ter VJ, De JT. Supporting reflection to improve learning from self-generated concept maps. Metacognition and Learning. 2022; 17(3):691-713.
[43] Bleckmann T, Friege G. Concept maps for formative assessment: creation and implementation of an automatic and intelligent evaluation method. Knowledge Management & E-Learning. 2023; 15(3):433-47.
[44] Göhler V, Yadav G. Analyzing concept maps in computer science education: an unsupervised learning approach with graph neural networks. In proceedings of the 25th annual conference on information technology education 2024 (pp. 56-61).
[45] Triantafyllou SA. Constructivist learning environments. In proceedings of the 5th international conference on advanced research in teaching and education 2022(pp.1-6).
[46] Cole AW, Lennon L, Weber NL. Student perceptions of online active learning practices and online learning climate predict online course engagement. Interactive Learning Environments. 2021; 29(5):866-80.
[47] Momen A, Ebrahimi M, Hassan AM. Importance and implications of theory of bloom's taxonomy in different fields of education. In international conference on emerging technologies and intelligent systems 2022 (pp. 515-25). Cham: Springer International Publishing.
[48] Stanton JD, Sebesta AJ, Dunlosky J. Fostering metacognition to support student learning and performance. CBE—Life Sciences Education. 2021; 20(2):1-7.
[49] Hristov S, Nakov D, Miočinović J. Constructive alignment between objectives, teaching and learning activities, student competencies and assessment methods in higher education. Journal of Agriculture and Plant Sciences. 2023; 21(2):21-36.

Sunu Mary Abraham is an academic professional with over 20 years of dedicated teaching experience at the postgraduate level. She is currently affiliated with Rajagiri College of Social Sciences, where she specializes in Design and Analysis of Algorithms, Machine Learning, and Data Analytics. She has successfully completed a research project funded by the Division of Educational Research (DER), NCERT, under the ERIC scheme. Presently, she is pursuing a Ph.D. in Learning Analytics at Avinashilingam Institute for Home Science and Higher Education for Women (Deemed to be University), with her research focused on advancing the field of Learning Analytics. Her work leverages data-driven methodologies to enhance the effectiveness of educational practices.
Email: sunumary@rajagiri.edu

G. Sudhamathy holds a Ph.D. in Computer Science from Bharathiar University, India. She is currently affiliated with Avinashilingam University and brings over a decade of IT industry experience alongside 15 years of academic expertise. Her research interests include Predictive Modeling, Web Log Analytics, and R Package Classifiers, reflected in her 15 journal publications and a published book. As an invited lecturer at AMA International University, Bahrain, she actively pursues professional development through workshops and refresher courses, highlighting her commitment to continuous learning. A versatile professional, she seamlessly integrates industry insight with a strong dedication to research and education.
Email: sudhamathy_cs@avinuty.ac.in

Appendix I

S. No.	Abbreviation	Description
1	AI	Artificial Intelligence
2	AISLE	Artificial Intelligence-Based Student Learning Evaluation
3	CM	Concept Map
4	FISCMAP	Fuzzy Based Inference System Model for Concept Maps
5	GNN	Graph Neural Network
6	GSKCM	Graphing Students' Knowledge using Concept Maps
7	JSON	JavaScript Object Notation
8	KLD	Kullback-Leibler Divergence
9	LO	Learning Outcome
10	LPG	Learning Path Generation
11	PHP	Hypertext Pre-Processor
12	UFS	Unsupervised Feature Selection
13	ML	Machine Learning