



Construction and Analysis of Collaborative Educational Networks based on Learner Concept Maps

HAYDEN FREEDMAN, NEIL YOUNG, DAVID SCHAEFER, QINGYU SONG,
ANDRÉ VAN DER HOEK, and BILL TOMLINSON, University of California, Irvine, USA

Network Analysis has traditionally been applied to analyzing interactions among learners in online learning platforms such as discussion boards. However, there are opportunities to bring Network Analysis to bear on networks representing learners' mental models of course material, rather than learner interactions. This paper describes the construction and analysis of collaborative educational networks based on concept maps created by undergraduates. Concept mapping activities were deployed throughout two separate quarters of a large General Education (GE) course about sustainability and technology at a large university on the West Coast of the United States. A variety of Network Analysis metrics are evaluated on their ability to predict an individual learner's understanding based on that learner's contributions to a network representing the collective understanding of all learners in the course. Several of the metrics significantly correlated with learner performance, especially those that compare an individual learner's conformity to the larger group's consensus. The novel network metrics based on collective networks of learner concept maps are shown to produce stronger and more reproducible correlations with learner performance than metrics traditionally used in the literature to evaluate concept maps. This paper thus demonstrates that Network Analysis in conjunction with collective networks of concept maps can provide insights into learners' conceptual understanding of course material.

CCS Concepts: • **Human-centered computing** → **Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: CSCL, Network Analysis, concept mapping, education

ACM Reference Format:

Hayden Freedman, Neil Young, David Schaefer, Qingyu Song, André van der Hoek, and Bill Tomlinson. 2024. Construction and Analysis of Collaborative Educational Networks based on Learner Concept Maps. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 36 (April 2024), 22 pages. <https://doi.org/10.1145/3637313>

1 INTRODUCTION

The Computer-Supported Collaborative Learning (CSCL) community has frequently turned to Network Analysis (often referred to as Social Network Analysis) to supplement traditional pedagogical assessments in characterizing learner understanding [33, 49, 56, 66]. Network Analysis in CSCL is typically performed over webs of nodes (usually representing individual learners), and edges (usually representing interactions or relationships between learners). The idea behind this kind of analysis is that individual actors are embedded in webs of relationships with other actors, and that these webs can be analyzed to gain insights about a particular actor based on that actor's ties to others [7]. CSCL seeks to understand the relationships between learning, learner interaction, and digital technologies; as such, Network Analysis is a natural fit for CSCL due to its ability

Authors' address: Hayden Freedman, hfreedma@uci.edu; Neil Young, neily@uci.edu; David Schaefer, drschae@uci.edu; Qingyu Song, qingys08@gmail.com; André van der Hoek, andre@ics.uci.edu; Bill Tomlinson, wmt@uci.edu, University of California, Irvine, Irvine, CA, USA, 92697.



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2024 Copyright held by the owner/author(s).

ACM 2573-0142/2024/4-ART36

<https://doi.org/10.1145/3637313>

to characterize these relationships in a quantifiable way [32, 47]. Various studies have provided evidence that analyzing learner interaction with Network Analysis techniques can provide insights such as identifying learner roles [55, 65], assessing learner problem-solving abilities [74], and understanding how learners make sense of complex systems [16]. It has been argued that using Network Analysis over such a network of learners and their interactions is key to understanding how learning occurs [18, 73].

However, CSCL has not fully taken advantage of the power of Network Analysis for investigating learner cognition. Cognitive Science research indicates that the cognitive processes in our brains form complex systems that help us solve problems, and that Network Analysis has enormous potential to model and investigate these processes [60, 61]. Unfortunately, current CSCL Network Analysis methodologies do not typically seek to construct or analyze networks of the cognitive models of learners. The current work takes a step towards advancing the subfield of Network Analysis in CSCL by taking advantage of insights gleaned from the application of Network Analysis in the field of Cognitive Science to model learner conceptions themselves as a network.

This study moves the conversation away from Social Network Analysis and towards a form of Epistemic Network Analysis [58] in which, rather than analyzing online interactions between learners in a discussion board setting, the focus is on analyzing learners' mental models about the central themes of a course. Learner concept map submissions are merged together to form a collaboratively-constructed collective network that can be seen as the entire class's consensus about the course material. Each individual learner map is then assessed based on its conformity to the course-wide consensus. This methodology provides several benefits, including 1) measuring what a learner knows at a given moment and how their knowledge changes over time, 2) understanding what the learners collectively know and how this collective understanding changes over time, and 3) understanding how an individual learner's mental model compares to the consensus of the larger group.

The work in this paper is novel in that it uses a group consensus-based approach to evaluate individual learners, based on the relative positioning of the learners' contributions within the collective network. The motivation for doing so is based on past results indicating that a group's collective mental model can approximate that of an expert [4, 26, 38]. The specific contributions of this work are 1) a novel methodology for collecting, merging, and analyzing concept maps generated by learners, and 2) an empirical comparison of a variety of novel and traditional metrics for evaluating learners concept maps using data from two offerings of an undergraduate course.

2 BACKGROUND AND RELATED WORK

2.1 Network Analysis in CSCL

Network Analysis research in CSCL settings has examined whether network metrics can be indicative of a learner's performance. Traditionally, this has been done by studying unimodal networks of learner interactions, often generated automatically from collaborative online learning platforms such as discussion boards [18, 56]. Unimodal networks are those which have only one type of network actor, such as learners. These networks are often constructed using digital trace data such as log files documenting online interactions amongst learners or between learners and teachers [2, 39, 53]. These works often argue that learners who engage more with other learners are likely to perform better, and this is tested by comparing the centrality of a learner's position within a network of learners to some external performance metric [8]. The majority of studies found strong correlations between the centrality of a learner in a network of interactions and learner performance [13, 51, 62], though at least one found little to no correlation [47].

Although creating and analyzing social interaction networks is convenient and accessible, focusing solely on this type of data limits the understanding that can be gleaned about learning patterns in CSCL. These types of learner interaction networks are generally, though not always, unimodal. A unimodal network of learners does not contain information on *what content* was discussed, only *who discussed it*; thus, it can be difficult to track learning about specific topics using this approach.

A small amount of past work has focused on networks connecting learners with course topics (i.e., bimodal networks). Agarwal and Ahmed [3] parse learner editing of Wiki pages to create a bimodal network of learners and pages, in order to assess learner collaboration and engagement. Likewise, Kim et al. [35] construct a bimodal network representing both which learners interacted with each other and which course topics they discussed. This approach has the advantage of providing information about more than just with whom learners interacted, additionally including information about which specific topics and themes from the course learners focused on.

A pair of literature surveys from the past decade have pointed out the lack of diversity of network actors in CSCL Network Analysis research. These surveys have made calls to expand the breadth of actors and types of relational ties used, as well as increase the diversity of the metrics used to analyze the networks and correlate the results with learner performance or learning outcomes [11, 18].

2.2 Concept Mapping

The phrase “concept mapping” is generally used to refer to Novakian concept mapping [9]. Novakian concept mapping is an educational activity in which learners create network-like representations of the relationships between related concepts [46]. For example, learners might map out their understanding of the Solar System by including concepts for planets, moons, and the Sun, and utilizing the relationship “revolves around” to connect the concepts in a logical manner.

Novakian concept mapping is often used to elicit an individual learner’s mental model within a particular domain and/or to track individual learning [19, 45, 48, 52]. Novakian concept maps benefit both learners and instructors. For instance, they help learners focus on the connections between course topics [15, 20, 63] as well as integrate knowledge across modules of a given course [14]. Additionally, they help instructors understand learner conceptions of course material, which is critical to improving teaching and achieving learning objectives [21, 40, 69].

Another technique commonly referred to as “concept mapping” is Group concept mapping. In contrast to Novakian concept mapping, Group concept mapping is a mixed-methods technique in which participants (often domain experts) brainstorm concepts related to a specific domain and sort these concepts into categories. Multidimensional Scaling (MDS) is applied to the participants’ submissions in order to automatically generate a map of concepts displayed in two-dimensional conceptual space [54]. In contrast to Novakian concept mapping, Group concept mapping does not seek to elicit the mental model of an individual nor generate relationships between concepts. It is instead concerned with the emergent properties of the algorithmically-generated group map, which may be used as a base of knowledge about the domain [54].

In this paper, learners constructed Novakian concept maps that are merged together to form a collective network. For clarity, future references in this paper to “concept maps” or “concept mapping” without a qualifier can be assumed to refer to Novakian concept maps.

Concept maps can be analyzed based on both content and structure. Often, the structural analysis of concept maps occurs in a qualitative manner; for example, maps can be visually classified using various structural templates such as “network”, “chain”, “tree”, etc. [5, 36]. It is also common to analyze maps via simple quantitative metrics such as the number of concepts or the number of relationships present in the map. More intricate concept map evaluation methods also exist; for example, Biswas et al. [6] deploy a teachable automated agent that takes a pre-programmed quiz

based on student concept map submissions; this allows students to receive real-time feedback on their maps.

2.2.1 Network Analysis and Concept Mapping. Both Group and Novakian concept maps are highly compatible with Network Analysis methodologies, due to their inherent network structure as well as their emphasis on the importance of relationships between entities. McLinden [44] notes that, while the goals of concept mapping and Network Analysis differ, the underlying data structure is the same. Although there are many studies exploring Network Analysis on Group concept maps [29, 44, 68, 71], there are far fewer exploring Network Analysis on Novakian concept maps. However, Network Analysis has the potential to quantitatively characterize individual Novakian concept map structure, which can inform the educator if the course goals are being met [60].

A common assumption underlying this claim is that the structure of a learner's concept map approximates the learner's level of understanding about a topic. Specifically, certain network structures are thought to be more indicative of expert-level knowledge, while others are more indicative of novice-level knowledge [38, 59]. Intuitively, those with more knowledge about a domain have a richer, more interconnected mental model than those with less knowledge about the domain.

In spite of the informative value of using Network Analysis with learner concept maps, only a small quantity of work has applied such techniques. In one such case, Siew et al. [59] analyze individual learner concept maps about Psychology using network metrics such as Average Shortest Path Length (ASPL) and Clustering Coefficient (CC), finding that these metrics were able to predict learner performance on quizzes. In another, Koponen and Nousiainen [38] create a collective network by merging 12 individual learner concept maps together, and use centrality metrics to identify key nodes in the network; however, they did not use the collective network to estimate individual learner performance. Schwendimann [57] uses a centrality metric to track changes in learner understanding of certain expert-determined "indicator concepts" via iterative concept mapping. The findings of Markham [43] show that maps about biological knowledge of mammals created by novices exhibited fewer hierarchical levels and fewer edges between concepts than maps created by experts.

2.2.2 Collaborative Concept Mapping. The subfield of Computer-Supported Collaborative Concept Mapping (CSCCM) defines collaborative concept mapping as two or more individuals using an online system to collaboratively work together on constructing one or more concept maps as a tool to facilitate shared understanding and construct knowledge [24, 27, 37, 42]. Studies have shown the effectiveness of this approach in facilitating problem solving [22, 64].

An alternative, less common approach to collaborative concept mapping involves learners working on concept maps individually, after which the individual maps are merged together to create a representation of the collective mental model of the participants [12, 17, 38]. A slight variation on this approach allows learners to optionally "share" specific elements from their individual map to the collective representation [50]. One benefit of this approach is that the collective representation is an aggregation of each individual's mental model on the topic. Thus, the contribution of each individual can be identified and analyzed.

2.2.3 Concept Mapping as Consensus Forming. Concept mapping activities with multiple participants can be viewed as a consensus forming activity, in which the prior ideas of individual learners are built upon via group communication [28] and the final map reflects the agreement of multiple mental models [25]. In some cases, a group consensus on concept map content and structure can be formed through social phenomena such as negotiation [41] or mediated via online learning tools [12]. In other cases, consensus is formed via merging individual learner maps without direct

social interaction between the participants. For instance, Koponen and Nousiainen [38] find that knowledge of physics concepts was highly dispersed among learners, but a collective network aggregating each individual learner’s map nearly perfectly matched a similar map created by an expert.

In this paper, we hypothesize that learners whose maps align more closely with the group consensus will have better understanding of the course material, and therefore higher performance. In order to test this hypothesis, several novel centrality and consensus-based metrics are compared against traditionally-used metrics, based on their ability to evaluate individual learner performance.

3 METHODS

3.1 Deployment of Assignments

Concept mapping assignments were deployed in two separate iterations of the same undergraduate course at a large US research university over the course of two academic quarters, enrolling a total of 679 participants. The course is a general education course about the intersection of sustainability and technology, and enrolls students from a wide variety of schools, majors, and academic standing within the university. Throughout both iterations, learners were asked to make concept maps by creating nodes (representing concepts) and linking them with edges (representing relationships). For example, learners might propose that “biodiversity increases sustainability.” By doing this repeatedly, learners created networks of related concepts and their relationships.

For both of the quarters, a dataset was generated containing the statements in each individual learner concept map along with a score for that map. Scores were assigned on a per-statement basis, either by a member of the research team or a team of Teaching Assistants (TAs). Statements were awarded a score of 1 if they were correct and relevant to the course material, and a score of 0 if they were incorrect or irrelevant to the course material. The learner’s final score was the percentage of statements they included in their concept map that received a score of 1.

In Spring ’21, learners used the freely available software CmapTools [10] to construct their concept maps, and in Fall ’21, learners used a custom concept mapping tool created by the research team. Figure 1 shows examples of two concept maps, one being constructed in CmapTools and the other in the custom tool.

The differences between the assignments across both quarters are summarized in Table 1.

Table 1. Quarter-by-Quarter Concept Map Assignment Details

Academic Quarter	Participating learners	Submission Time (10 Week Quarters)	Node Label Set	Edge Label Set
Spring 2021	139	Week 10	All English Wikipedia articles	20 labels selected by instructor only "causes"
Fall 2021	251	Week 5	All English Wikipedia articles	
Academic Quarter	Submission Rounds		Scoring System	
Spring 2021	3		Single member of research team scores all relationships	
Fall 2021	1		Team of TA's scores all relationships after performing consensus building activity	

3.1.1 Spring ’21. The first deployment of the concept mapping assignment occurred in the Spring ’21 iteration of the course. To ensure consistency and avoid the use of synonyms between maps, the assignment restricted learners to use concepts from the list of all English Wikipedia article titles and relationships from a list of 20 provided by the instruction staff. Learners were also required to include the two central themes of the course, “sustainability” and “technology”, as nodes in their maps.

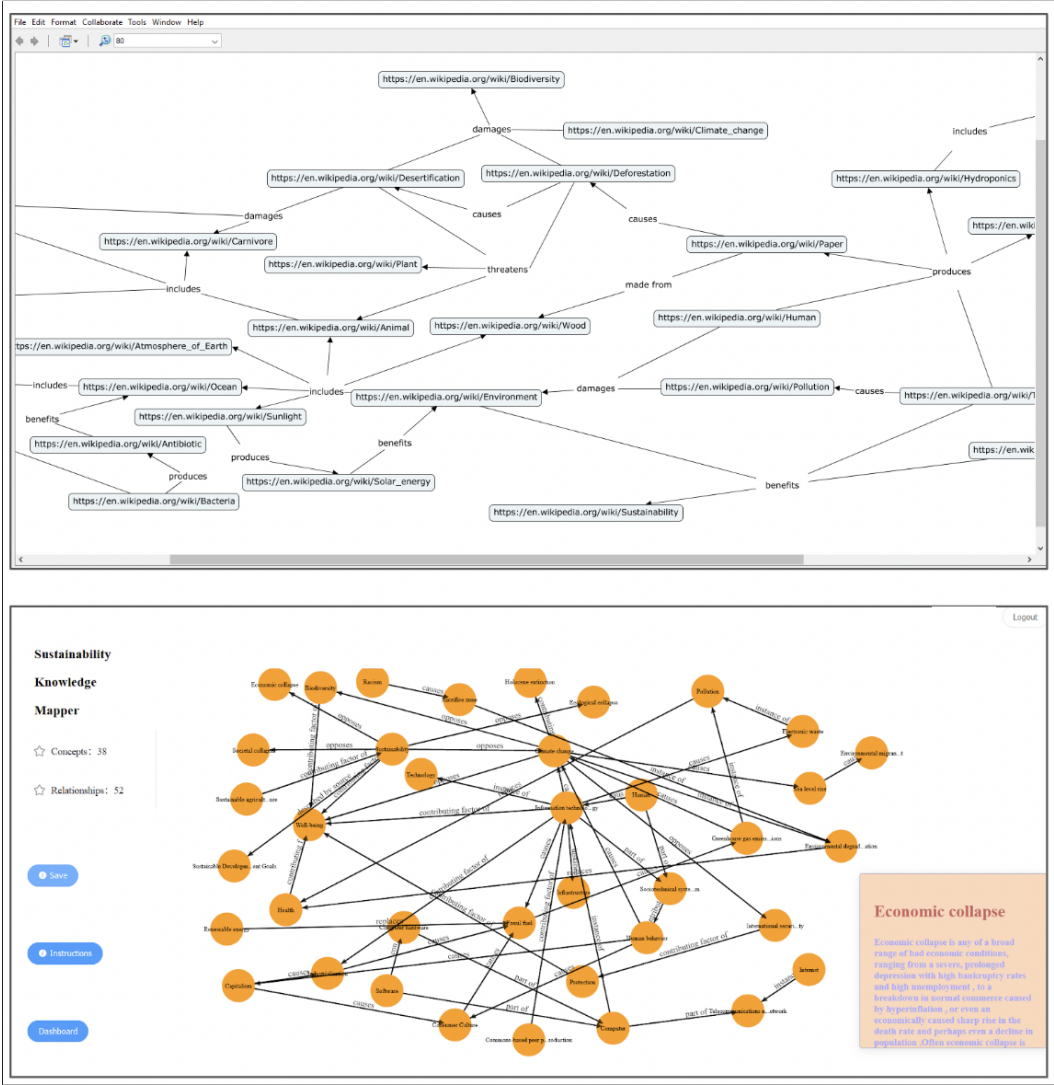


Fig. 1. The upper screenshot shows part of one learner’s concept map submission from the Spring ’21 quarter in the CmapTools interface. The lower screenshot shows an example concept map created by a research team member using the custom concept mapping interface used for the Fall ’21 quarter.

This quarter focused on learners improving their maps iteratively. In the first assignment, learners created an initial map, while in the second and third assignments, learners revised their maps based on course material introduced since the previous revision. The data analyzed for this study was taken from the third and final concept mapping submission. A member of the research team assigned scores to each statement, which were used to calculate the final per-learner scores.

In order to better focus the analysis on learners’ original contributions, the two required nodes ("sustainability" and "technology") were removed from the collective network before the Network Analysis metrics were calculated.

3.1.2 Fall '21. In the second deployment, learners were again required to use Wikipedia articles as concepts. However, as the instructor wanted to examine the ways that learners constructed causality networks, the only allowed relationship was “causes”. Learners were not required to include any specific nodes in their submissions; thus, this quarter’s assignment was slightly less restrictive in terms of possible concepts but much more restrictive in terms of permitted relationships.

In this quarter, there was no revise-and-resubmit process as in the previous quarter; the data analyzed is from the learners’ first and only submissions. Learners submitted the assignments midway through the quarter. Each individual statement was assigned a score of 0 or 1 via a TA review process, in which the TAs first participated in a consensus forming activity to establish shared criteria and then were assigned an anonymized spreadsheet containing a portion of the learner statements to review and score.

3.2 Merging of Individual Concept Maps into a Collective Network

By merging individual learner concept maps together, one can form a network that exhibits the class’s collective understanding of the material. However, in order to create a collective network, a merge strategy must first be selected.

Such a strategy must define how duplicate edges between individual maps will be handled (i.e., whether to weight an edge by the number of learners reporting it or represent all edges as unweighted). A weighted strategy places higher values on edges added by more students, while an unweighted strategy allows for the possibility that individual students made unique but valuable edges not included by other students.

A merge strategy must also define whether the directionality of the relationships will be preserved. For example, say that a learner argues that Climate Change → causes → Sea Level Rise. A directed merge strategy would preserve the fact that Climate Change “points” at Sea Level Rise. An undirected merge strategy would omit this information, and merely note that these two nodes were connected by at least one learner.

A directed merge strategy preserves the most information from the original concept map; however, an undirected merge strategy acknowledges that the semantics of the relationship label affect the directionality in an arbitrary way. For example, if the relationship label “causes” were to be replaced by “caused by”, the directionality of every relationship using this label would be reversed, but this says nothing about the actual relationship between the two nodes, only the semantics. By using an undirected strategy, we remove the bias introduced by the semantics of the relationship label.

When creating the collective networks, only the presence or non-presence of an edge between two nodes was included; the specific label chosen by learners for each relationship was ignored. This was done in order to allow for more overlap between individual learner maps in the collective network, and because, from a network perspective, the structure of the connections learners made between concepts was more interesting than the specific semantics they used to make the connections.

For example, if one learner included the statement Climate Change → synonym → Global Warming, and another learner included the statement Climate Change → causes → Global Warming in their respective maps, the collective map would contain one of four possible edges based on the merge strategy selected: 1) Climate Change → Global Warming with a weight of 2 for the directed/weighted strategy, 2) Climate Change → Global Warming with no weight for the directed/unweighted strategy, 3) Climate Change – Global Warming with a weight of 2 for the undirected/weighted strategy, and 4) Climate Change – Global Warming with no weight for the undirected/unweighted strategy. Here, “→” implies a directed edge and “–” implies an undirected edge. Figure 2 shows the four types of merge strategies used in the analysis.

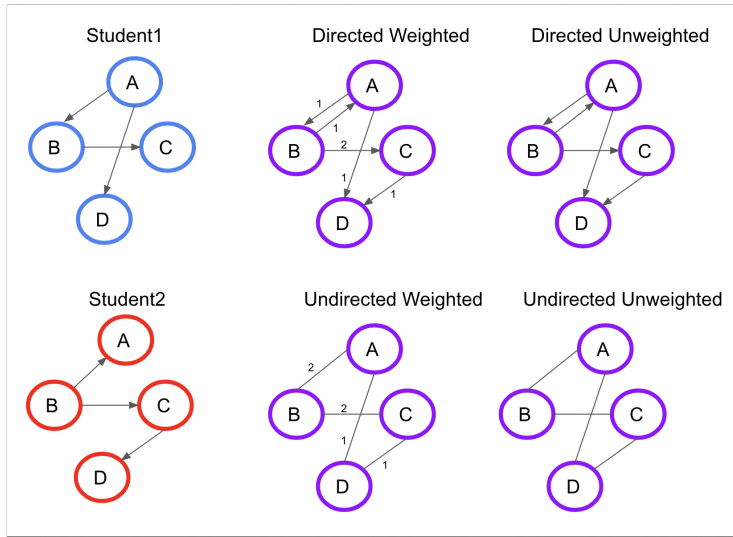


Fig. 2. Two example learner concept maps merged into a collective network using each of the 4 merge strategies used to conduct the analysis. The networks with blue (top left) and red (bottom left) nodes represent each of the two learner maps, and the networks with purple nodes (four most rightward) represent the 4 possible collective networks.

Due to differences in assignment requirements and scoring between the two quarters, combining both quarters of data into a single collective network was problematic. Instead, individual collective networks were created for each quarter.

3.3 Analysis of Maps

Analysis of the learner concept maps is broken down into several categories:

- Qualitative metrics*
- Individual concept map metrics*
- Centrality metrics**
- Consensus metrics**

Categories annotated with a single asterisk (*) are traditional Network Analysis metrics. Categories annotated with a double asterisk (**) are novel metrics using the methodology of merging individual learner concept maps into a collective network. The NetworkX Python library [31] was used to carry out the analyses, aside from the qualitative metrics. Note that only Comprehensiveness (in the qualitative metrics category) considers the actual content of the concept map; all of the other metrics are based either on individual map structure (traditional metrics) or the positioning of individual learner maps within the collective network (novel metrics).

3.3.1 Qualitative Metrics. Qualitative metrics (also referred to as holistic metrics) are one of the most common forms of concept map evaluation. Evaluation of such metrics is typically carried out by an instructor or expert looking at a learner's concept map and assessing its merit based on content and/or structure, but without performing any counting or computation. Besterfield-Sacre et al. [5] define a rubric for several qualitative concept map metrics, including Comprehensiveness (the degree to which a map's content covers the relevant material) and Organization (orderliness of the arrangement of nodes and edges). Additionally, Yin et al. [72] define five common structural

templates (Structural Form) that can be used to classify the structures of learner concept maps: "linear", "circular", "hub-spoke", "tree", and "network". The "network" structural template is characterized by a web of interconnected concepts, and is considered more indicative of meaningful learning than the other structural templates [36].

Table 2. Summary Descriptions of Qualitative Metrics

Individual Metric	Measure	Range of Output
Comprehensiveness	Completeness of content	Score of 1, 2, or 3, where higher number is more comprehensive
Organization	Orderliness of map arrangement	Score of 1, 2, or 3, where higher number is more organized
Structural Form	Structural form of concept map	One of "linear","circular","hub-spoke","tree","network"

For the qualitative metrics, one member of the research team reviewed every individual learner concept map from both academic quarters and assigned each a score for Comprehensiveness, Organization, and Structural Form. Table 2 describes the qualitative metrics used in this analysis. One issue with qualitative analysis of concept maps is that such analysis is subjective in nature. For instance, many maps share characteristics of two or more of the structural templates, leaving it up to the scorer to make a judgment call.

3.3.2 Individual concept map metrics. The quantitative analysis of the individual learner concept maps includes both simple counting-based metrics such as number of concepts and number of relationships, as well as traditional network metrics for concept map analysis: Average Shortest Path Length (ASPL), Clustering Coefficient (CC), Network Density, and Complexity. ASPL and CC are metrics that can be used to identify Small-World Networks (SWNs). SWNs are characterized by highly clustered groups and short path lengths from any given node to another. These types of networks are representative of many types of real-world systems [70]. Past research has hypothesized that learners who build concept maps that conform to the characteristics of SWNs have a more complete understanding of the course material. Specifically, Siew [61] found that learners with lower ASPL scores and higher CC scores performed better on quizzes.

Figure 3 shows the structure of the individual learner concept maps with the highest and lowest scores for the metrics ASPL, CC, and Density from the Fall '21 quarter. This figure demonstrates that such metrics can be used as a quantitative alternative to the typically qualitative task of characterizing concept map structure. For example, both high ASPL and low CC scores are associated with what would qualitatively be labeled as a "chain" structure, whereas low ASPL and high Density scores reflect more of a "network" structure in the qualitative coding.

Table 3. Summary Descriptions of Individual Network Metrics

Individual Metric	Measure	Interpretation of Output
Number of Concepts	Size of network nodeset	More concepts is larger nodeset
Number of Relationships	Size of network edgeset	More relationships is larger edgeset
ASPL	Interconnectedness of network	Lower number is more interconnected
CC	Extent to which nodes cluster together	Higher number is more clustered
Network Density	Number of existent edges compared to all possible edges	Higher number is denser
Complexity	Ratio of edges to nodes	Higher number is more complex
Number of Hierarchies	Degree to which network is organized into branches	More hierarchies is more organized into branches
Level of Deepest Hierarchy	Depth of knowledge	Higher level is deeper knowledge
Number of Cross-links	Degree to which hierarchies are connected	More cross-links is more connected hierarchies

Another category of individual concept map metric is based on identifying hierarchy within a concept map. Besterfield-Sacre et al. [5] define a hierarchy as a shortest path from a root concept (a

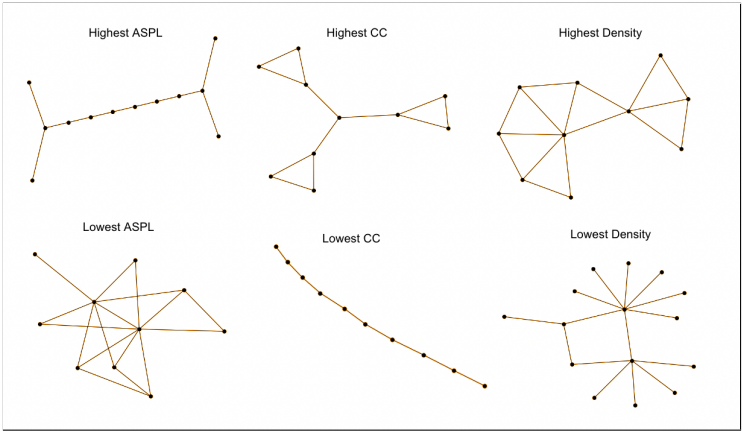


Fig. 3. The learner networks with the highest and lowest scores for three of the individual network metrics from the Fall ’21 quarter, showing a diversity of network structures in individual learner submissions

concept with no parents) to a leaf concept (a concept with no children). They define three hierarchy-based metrics that are used for quantifying concept map structure: number of hierarchies in the map, length of the deepest hierarchy, and number of cross-links between hierarchies. A higher total number of hierarchies, deeper hierarchies, and higher number of cross-links are associated with more complex concept map structure.

Table 3 contains descriptions of each of the individual metrics.

3.3.3 Centrality metrics. The centrality-based metrics measure whether learners are able to identify key nodes, as determined by their peers. Individual learner scores are calculated based on the average centrality of the nodes in the collective network that were part of their individual concept map. Three centrality metrics were calculated: Betweenness, Degree, and Closeness (described in Table 4).

Table 4. Summary Descriptions of Node Centrality Network Metrics

Centrality Metric	Measures	Interpretation of Output
Node Betweenness	Number of shortest paths that pass through a node	Higher number is more central
Node Closeness	Reciprocal of sum of shortest paths between a node and all other nodes	Higher number is more central
Node Degree	Number of edges connected to a node	Higher number is more central

Figure 4 shows the positioning of two individual learner networks (red edges) within the collective network of all the learner concept maps from the Fall ’21 quarter (blue edges). While one learner’s individual concept map is positioned in the center of the collective network (left), the other learner’s map passes through the center but mostly lies on the periphery of the collective network (right). We hypothesized that learners who made maps that are positioned more centrally within the collective network will have higher scores on the concept map assignment.

3.3.4 Consensus metrics. While the centrality metrics measure learner identification of key nodes, this section explores four metrics that instead measure learner identification of the correct edge structure between nodes. Specifically, these metrics compare an individual learner contribution to the local region of the collective network pertaining to the same nodeset (i.e., the set of concepts

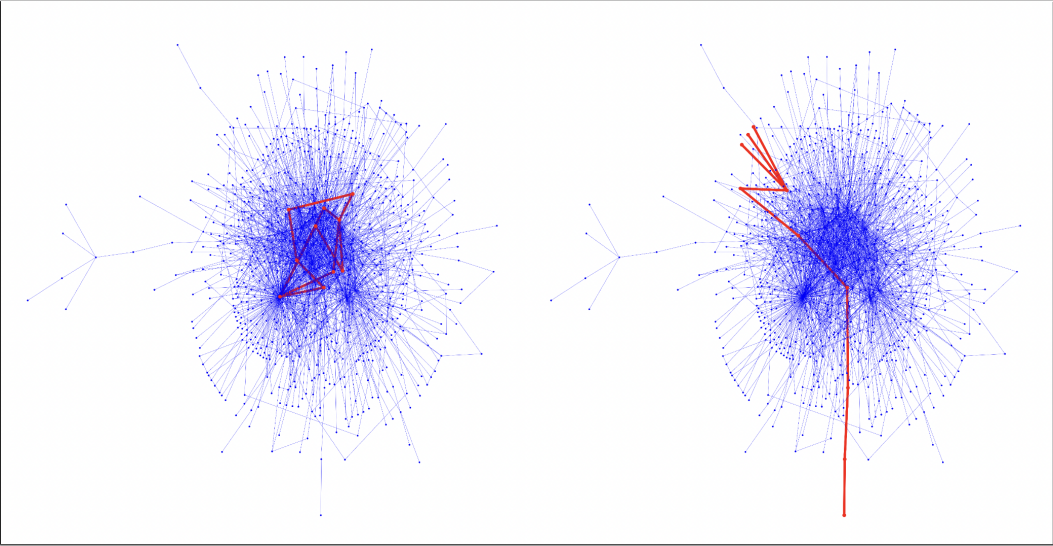


Fig. 4. Two individual learner concept maps from Fall 2021 layered in red over the course-wide collective network in blue. The learner map on the left exhibits a relatively high Node Betweenness score, whereas the map on the right exhibits a relatively low Node Betweenness score.

included in the individual learner’s concept map). Thus, they measure the "edge overlap" that an individual learner concept map shares with the maps of other learners.

These metrics are important because global centrality metrics that consider the entire collective network do not account for how a learner’s choice of nodeset may limit their centrality scores. Consider a learner who makes a high quality concept map but leaves out a single key node that was referenced by the majority of the other learners. This would substantially reduce their centrality score, which is based on which nodes a learner included. The consensus-based metrics introduced in this section do not directly penalize a learner for omitting popular nodes. Instead, they assess the ways that other learners collectively connected nodes within the same local network region, allowing for learners who chose to include a more diverse nodeset to still potentially receive high scores.

Table 5. Summary Descriptions of the Consensus Network Metrics

Consensus Metric	Measures	Interpretation of Output
Edge Consensus	# of overlapping connections made by other learners	Higher number is more overlapping
Subgraph Coverage	# of "missed" edges made by other learners for a given set of nodes	Lower number is more overlapping
Collective Shortest Path	Average shortest alternative path between pairs of nodes	Lower number is more overlapping
Communicability	Weighted sum of all paths between pairs of nodes	Higher number is more overlapping

Table 5 shows descriptions of each of the four consensus metrics.

Edge Consensus

Edge Consensus assesses how well a learner conforms with the general agreement of the class based on whether pairs of nodes are connected via edges. Each edge in the concept map of an individual learner is assigned a score based on the number of other learners that also included that edge in their map. The final score is calculated from the summed score of all of the edges in their

map, divided by the total number of edges they included. The simplicity of this metric makes it easy to calculate, even on large networks.

We hypothesized that Edge Consensus will positively correlate with individual learner map quality, as it rewards learners who share a common understanding of the course material with their classmates. While the metric does not directly reward learners for including popular nodes, it does reward learners for including popular edges, which will indirectly lead to learners who included popular nodes receiving higher Edge Consensus scores.

Subgraph Coverage

Calculating Subgraph Coverage involves creating a subgraph of the collective network defined as the set of edges added by all of the learners between the nodes referenced in an individual learner's map. The number of edges in the individual map is then compared with the number of total edges in the subgraph. The intuition behind this metric is to evaluate the edge coverage that an individual learner has made over their selected nodeset compared to the rest of the class.

It is hypothesized that higher Subgraph Coverage will be associated with higher individual concept map quality. In theory, learners who were able to cover more of their selected subgraph will be demonstrating a more thorough understanding of the concepts in the subgraph than learners who covered less of the subgraph.

Collective Shortest Path

The Collective Shortest Path metric assesses the ways that other learners collectively represented the relationships present in each individual learner's concept map. For each individual concept map, a modified collective network is created based on all of the other learners' concept maps but excluding the current learner's map. Then, for each edge connecting a pair of nodes in the learner's map, the shortest path between this pair of nodes in the modified collective network is calculated. If another learner added the same edge, then the shortest path is 1; if a path does not exist, then the longest shortest path in the collective network is assigned as the score for the edge in question.

The Collective Shortest Path metric rewards learners who included edges also included by other learners, which is also the case in the Edge Consensus metric. However, Collective Shortest Path also rewards learners who made unique connections between nodes that are in adjacent regions of the collective network, even if no other learner directly connected those nodes. Conversely, it punishes learners that connected nodes that would otherwise have been far away from each other in the collective network.

The guiding principle behind this metric is that learners may include unique edges that are still of high quality; in fact, unique edges may actually provide an important bridge between concepts that would otherwise not have been connected. However, we argue that unique edges of high quality are more likely to occur between two nodes that are already close together, whereas unique edges of low quality may bridge two nodes that otherwise would have been far apart. Following from this, we hypothesized that learners who receive a lower Collective Shortest Path score will have higher quality concept maps.

Communicability

Unlike the other three metrics in this section, Communicability is borrowed from previous literature [23]; however, its application to assigning individual learner scores based on their contributions to a collective network is novel. It uses the same modified collective network as Collective Shortest Path; however, in contrast to Collective Shortest Path, which only considers the shortest alternative

path between each pair of nodes, Communicability applies a weighted sum of the lengths of all alternative paths in the collective network between every pair of nodes in each individual student network, such that longer paths are weighted less heavily than shorter paths. It rewards learners that connected nodes that were reachable by a wide variety of other paths, indicating that such learners made connections in important parts of the collective network. Therefore, we hypothesized that a higher score will lead to higher quality concept maps.

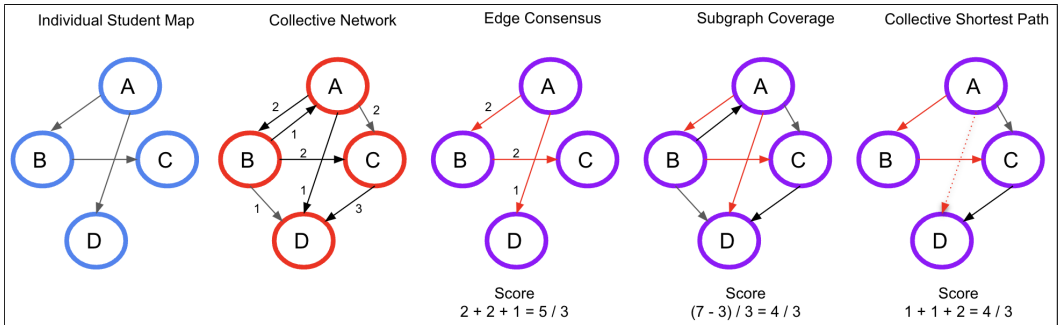


Fig. 5. An example individual learner map (left, blue), alongside a collective network that includes the individual learner map (second from left, red), and the calculation of three novel consensus metrics (three most rightward, purple). In the three purple graphs showing the calculations, red arrows indicate edges in the individual learner map, black arrows indicate edges in the collective network not present in the individual learner map, and the dotted red arrow indicates an edge that is present in the individual learner map that was not made by any other learner. These examples use the weighted/directed merge strategy. For Edge Consensus, the weights of the relevant edges in the collective network are summed. For Subgraph Coverage, the individual learner map contains 3 of the 7 edges (highlighted in red) that the collective network contains between the nodeset referenced in the individual learner map. For Collective Shortest Path, 2 of the 3 edges in the individual learner map were referenced by at least one other learner, leading to scores of 1 for those edges. The final edge is assigned a score of 2 as the algorithm found a shortest path from A→D through the node C.

Figure 5 is a visualization showing a simple calculation of each of the three consensus metrics introduced in this paper (excluding Communicability, which is detailed in [23]).

4 RESULTS AND DISCUSSION OF RESULTS

4.1 Summary Statistics

Summary statistics for the collective networks are shown in Table 6. Learners in the Spring '21 quarter made larger networks than in the Fall '21 quarter, in terms of both nodes and edges. This is likely due to differences in the assignment instructions; learners had one edge label to choose from in the fall, vs. 20 in the spring. However, Network Density and Average Node Degree are similar between the networks, indicating that the structure of the two collective networks is similar. The average learner score differs greatly between quarters due to different scoring methodologies; the TA grading team assigned much higher learner map scores than the research team member, leading to a much higher average score for the Fall '21 quarter than the Spring '21 quarter.

Table 7 shows summary statistics for the qualitative metrics. Learner maps in the Spring '21 quarter scored higher for both Comprehensiveness and Organization than maps for the Fall '21 quarter. The Spring '21 quarter data also contained far more learner maps coded as a "network" structure than Fall '21 (88.5% of all maps vs 27.9% of all maps). Summary statistics for quantitative metrics are shown in Table 8. The mean, standard deviation, and range of each metric output are shown.

Table 6. Summary Statistics for Quarter-by-Quarter Collective Networks. Here, the “Number of Statements” column is the total number of statements in all of the individual learner submissions, whereas the two “Number of Edges” columns show the number of edges remaining in the collective network once directed or undirected merge strategies have been applied.

Academic Quarter	Number of Nodes	Number of Statements	Number of Edges - Directed Merge	Number of Edges - Undirected Merge	Mean (SD) of Learner Score	Network Density	Average Node Degree
Spring '21	1,145	4,903	3,683	3,422	.31 (.14)	0.005	5.98
Fall '21	724	2,605	2,001	1,908	.76 (.19)	0.007	5.27

Table 7. Summary Statistics for Qualitative Metrics

Metric	Spring '21	Fall '21
Comprehensiveness	1 : 40%, 2 : 50%, 3 : 10%	1 : 49%, 2 : 39%, 3 : 11%
Organization	1 : 24%, 2 : 52%, 3 : 24%	1 : 63%, 2 : 26%, 3 : 11%
Hub-spoke	1.40%	8.30%
Tree	5.80%	31.40%
Linear	3.50%	24.70%
Network	88.50%	27.90%
Circular	0.70%	7.60%

Table 8. Summary Statistics for Quantitative Metrics. The Mean, Standard Deviation (SD), and Range (RNG) of all metric outputs are shown, broken down by quarter. The Mean is represented as the first number in each cell, while the SD and RNG are shown in parenthesis.

Metric	Spring '21	Fall '21
ASPL	1.01 (SD = .79, RNG = 4.34)	.81 (SD = .64, RNG = 4.91)
CC	.09 (SD = .08, RNG = .32)	.04 (SD = .07, RNG = .35)
Density	.08 (SD = .04, RNG = .19)	.11 (SD = .02, RNG = .12)
Complexity	.70 (SD = .16, RNG = .67)	.99 (SD = .13, RNG = .58)
Number of Concepts	24 (SD = 7, RNG = 27)	10 ¹ (SD = .6, RNG = 8)
Number of Edges	35 (SD = 9, RNG = 56)	10 (SD = 1.7, RNG = 11)
Number of Hierarchical Levels	8 (SD = 3, RNG = 12)	5 (SD = 2, RNG = 10)
Number of Hierarchies	15 (SD = 15, RNG = 91)	4 (SD = 3, RNG = 24)
Number of Cross-Links	17 (SD = 12, RNG = 45)	3 (SD = 3, RNG = 14)
Node Betweenness	.01 (SD = .005, RNG = .03)	.03 (SD = .02, RNG = .07)
Node Degree	.02 (SD = .009, RNG = .04)	.05 (SD = .02, RNG = .09)
Node Closeness	.16 (SD = .03, RNG = .17)	.21 (SD = .03, RNG = .21)
Edge Consensus	3.36 (SD = 1.79, RNG = 7.83)	2.41 (SD = 2.23, RNG = 11.10)
Subgraph Coverage	1.49 (SD = .76, RNG = 3.76)	.78 (SD = .51, RNG = 2.33)
Collective Shortest Path	5.11 (SD = 2.00, RNG = 9.53)	5.55 (SD = 3.05, RNG = 12.90)
Communicability	2.7X10 ⁸ (SD = 1.8X10 ⁸ , RNG = 8.6X10 ⁸)	2.7X10 ⁷ (SD = 1.7X10 ⁷ , RNG = 7.0X10 ⁷)

4.2 Correlation Results

To evaluate the performance of each metric, the statistical correlation between each network metric and the concept map score is calculated on a learner-by-learner basis. The Pearson correlation coefficient was used, except in the cases of the qualitative network structure metric, which is a categorical variable. To handle the statistical challenge of regression over a categorical and a continuous variable, this metric was broken out into five separate Boolean-valued metrics representing each of the five structural templates (i.e., "tree", "chain", etc.), and the point-biserial correlation was used.

In the tables below, each cell contains the Pearson or point-biserial correlation coefficient representing the correlation between the network metric and the learner concept map scores, as well as the p-value of the correlation. Due to the large number of statistical tests run in this analysis,

we apply Sidak-Holm adjustment to account for the Family-Wise Error Rate [30]. Based on this result, we consider $p < .0006$ to be a statistically significant result. Such results are shown in **bold**. Results meeting the standard threshold of statistical significance of $p < .05$ are marked with an asterisk (*).

4.2.1 Traditional Metrics from the Literature.

Qualitative Metrics: The qualitative results in Table 9 show that, in spite of their frequent use in past literature [34], the qualitative metrics are not highly correlated with learner performance on the concept mapping assignment, with the exception of Comprehensiveness for the Fall '21 quarter, which addresses the scope of the content contained within the map rather than the structure of the map.

Table 9. Qualitative Metrics Results

	Comprehensiveness	Organization	Hub-spoke	Tree	Linear	Network	Circular
Spring '21	.13, $p = .116$	-.05, $p = .529$	-.05, $p = .09$	-.03, $p = .75$	.05, $p = .57$	.04, $p = .65$	.03, $p = .73$
Fall '21	.49, $p < .0006^*$	-.05, $p = .391$	.06, $p = .32$	-.08, $p = .21$	.10, $p = .11$	-.08, $p = .20$	.05, $p = .43$

Individual, Non-Hierarchy-based Metrics: The results of the quantitative metrics in Table 10 show that the individual, non-hierarchy-based metrics correlated with learner performance for the Spring '21 quarter but not for the Fall '21 quarter. Unexpectedly, CC and Density both correlated negatively with learner performance, where positive correlations were hypothesized in both cases. This can be explained by the observation that, for the dataset in question, CC and Density are both positively correlated with each other ($p < .0001$) and both are negatively correlated with number of concepts ($p < .0001$), while number of concepts was strongly positively correlated with learner performance. The general thinking behind using CC and Density to analyze concept maps is that they provide an easy-to-compute method to characterize the complexity of the network; however, a confounding factor is that a map with fewer nodes can lead to higher CC and Density scores, but also be indicative of lower learner effort.

Table 10. Individual, Non-Hierarchy-based Metrics Results. Note that Number of Concepts is not listed for Fall '21 because learners were required to include a fixed number of nodes for this quarter.

Academic Quarter	ASPL	CC	Density	Complexity	Num. Concepts	Num. Edges
Spring '21	-.25, $p = .004^*$	-.24, $p = .004^*$	-.36, $p < .0006^*$	.22, $p = .009^*$	.40, $p < .0006^*$	.22, $p = .01^*$
Fall '21	.07, $p = .24$	-.12, $p = .06$	-.11, $p = .07$	.08, $p = .19$	N/A	-.06, $p = .38$

Individual, Hierarchy-based Metrics: The individual, hierarchy-based results in Table 11 show that, while some of these metrics correlate with learner performance, none of the results are reproduced across quarters. The number of hierarchical levels was a useful predictor of performance in the Fall '21 quarter, and the number of hierarchies was highly correlated with performance for Spring '21.

Table 11. Individual, Hierarchy-based Metrics Results

	Num. Hierarchical Levels	Num. Hierarchies	Num. Cross-links
Spring '21	-.11, $p = .18$	.30, $p < .0006^*$	-.20, $p = .02^*$
Fall '21	.17, $p = .006^*$	-.04, $p = .49$	-.07, $p = .27$

In general, the traditional metrics from the literature do not excel at predicting learner performance across the two datasets used in this study. While there are some isolated statistically significant correlations, none of the results for any of the traditional metric categories were reproduced across both quarters of data. These results motivate the investigation of metrics that move beyond assessing individual learner conceptions in isolation to assessing them in relation to the conceptions of their peers.

4.2.2 Metrics based on Collective Network.

For the metrics in this subsection, there are four rows for each academic quarter, representing the results of the various merge strategies: unweighted/directed (UD), weighted/directed (WD), unweighted/undirected (UU), and weighted/undirected (WU).

Centrality Metrics: Table 12 shows the results of the centrality metrics. In general, the centrality metrics produced stronger correlations for the Fall '21 quarter, where the majority of the results were statistically significant.

Table 12. Centrality Metrics Results

Academic Quarter	Node Betweenness	Node Degree	Node Closeness
Spring '21 (UD)	.10, $p = .24$	.18, $p = .03^*$	.25, $p = .003^*$
Spring '21 (WD)	.10, $p = .23$	N/A	.23, $p = .01^*$
Spring '21 (UU)	.16, $p = .05$	.19, $p = .02^*$	.10, $p = .23$
Spring '21 (WU)	.17, $p = .05$	N/A	.08, $p = .37$
Fall '21 (UD)	.22, $p < .0006^*$	.28, $p < .0006^*$	.17, $p = .009^*$
Fall '21 (WD)	.21, $p < .0006^*$	N/A	.12, $p = .06$
Fall '21 (UU)	.24, $p = .001^*$	.28, $p < .0006^*$	.25, $p < .0006^*$
Fall '21 (WU)	.23, $p < .0006^*$	N/A	.20, $p = .002^*$

Consensus Metrics: Table 13 shows the results of the consensus metrics. These metrics produced a large number of statistically significant correlations with learner performance, including many significant results. As hypothesized, Edge Consensus and Communicability produced positive correlations with student score across both datasets, while Collective Shortest Path produced negative correlations. Interestingly, Subgraph Coverage produced strong positive correlations, despite negative correlations being hypothesized. In this metric, a lower score indicates higher coverage, and higher coverage implies more overlap with other learners. This means that concept maps with higher coverage contained lower quality content, an unexpected result.

However, an important downside of the Subgraph Coverage metric is that it has the potential to penalize learners that included popular nodes. Popular nodes are more likely to have many edges between them, so it is more difficult for an individual learner to cover all of these edges. Conversely, in the extreme case where a learner creates a concept map with a nodeset completely distinct from the nodes included by all other learners, this learner would achieve a perfect coverage score. This introduces a confounding factor in the Subgraph Coverage metric, indicating that the metric needs to be refined before it can be considered useful.

Table 13. Consensus Metrics Results. Note that only weighted merge strategies were applied to Edge Consensus, as this metric can only function in the presence of an edge weight. Also note that only the unweighted, undirected strategy was applied to Communicability, due to the NetworkX implementation of this metric.

Academic Quarter	Edge Consensus	Subgraph Coverage	Collective Shortest Path	Communicability
Spring '21 (UD)	N/A	.35, $p < .0006^*$	$-.25, p = .003^*$	N/A
Spring '21 (WD)	$.17, p = .04^*$	$.21, p = .01^*$	$-.27, p = .001^*$	N/A
Spring '21 (UU)	N/A	.38, $p < .0006^*$	$-.22, p = .009^*$	$.27, p = .001^*$
Spring '21 (WU)	$.20, p = .02^*$	$.24, p = .005^*$	$-.24, p = .005^*$	N/A
Fall '21 (UD)	N/A	.24, $p < .0006^*$	$-.39, p < .0006^*$	N/A
Fall '21 (WD)	.42, $p < .0006^*$	.21, $p < .0006^*$	$-.32, p < .0006^*$	N/A
Fall '21 (UU)	N/A	.21, $p < .0006^*$	$-.36, p < .0006^*$	$.32, p < .0006^*$
Fall '21 (WU)	.42, $p < .0006^*$	$.21, p = .001^*$	$-.31, p < .0006^*$	N/A

5 GENERAL DISCUSSION

This paper presents several novel Network Analysis metrics for analyzing learner concept maps based on merging individual maps into a collective network. Correlations are computed between centrality and consensus-based metrics derived from this collective network as well as from several categories of metrics traditionally used in the concept mapping literature.

On the dataset used in this paper, most of the novel metrics were better indicators of student performance than the traditional metrics. The results from the novel metrics were also more consistent across both academic quarters than the results from the traditional metrics. The consensus-based metrics were overall the best predictors of student map quality, although the centrality metrics were also relatively useful predictors. In contrast, the qualitative metrics and individual concept map metrics were less reliable predictors.

The simplicity of the qualitative and individual map metrics may partially explain their past use for concept map analysis tasks; however, these metrics consider each student's concept map individually and do not consider the interplay between an individual student's conceptions and the course's collective conceptions of the material. Creating a collective network of merged student concept maps enables a variety of novel metrics that treat the entire course's collective understanding as a proxy for expert understanding, providing a knowledge base against which individual student contributions can be compared.

An additional distinction between the traditional and novel metrics is that the traditional metrics typically make the assumption that a learner's performance is tied to the structural shape of their concept map, irrespective of the selected topics. In contrast, the novel metrics assign a score based on the positioning within the collective network of the concepts and relationships that a student included, without considering the structure of the individual student map. The results from the novel metrics indicate that a learner's correct selection of key nodes (centrality) as well as correct selection of specific edges between them (consensus) should be considered valuable indicators of learner performance.

While the novel metrics in this paper introduce a new framework for concept map evaluation, they are not perfect indicators of student performance and are not intended to replace traditional metrics. Past research has shown that many of the traditional metrics also provide predictive value. Particularly, it is widely agreed that concept maps demonstrating expert-level understanding more commonly assume a "network"-like structure with high interconnectivity between concepts, which can be assessed via qualitative evaluation or traditional network metrics such as ASPL. This type of evaluation is not enabled by the novel metrics, as these metrics only assess individual structure of the map in terms of its relation to the collective network.

Rather than replacing traditional metrics, the novel metrics introduced in this paper should be viewed as viable accompaniments to previous methodology. In many cases, the best results may be obtained from an ensemble of different metrics. As our own results show, the efficacy of various metrics depends largely on the data itself; however, practically no studies have tried to apply these metrics to a diverse set of concept map data from various educational contexts. Such a study is needed to better determine the generalizability of these metrics.

5.1 Future Work

This paper is in part a response to calls from the CSCL community to explore more diverse forms of network actors, relational ties, and metrics within the Network Analysis space. This call is answered via a Network Analysis of learner concept map data, a nascent but intriguing area of investigation. While initial results are promising, this work can be improved or extended in a few key areas.

For one thing, both of the datasets used for this study had different characteristics, largely due to the instructor varying the assignment parameters between quarters. These differences caused changes in the effectiveness of the metrics, especially the individual concept map metrics. This observation points to the importance of aligning one's metrics with respect to particulars of the concept mapping assignment parameters. As concept mapping assignments vary widely, further research is needed to pinpoint the most appropriate set of metrics for each variation or category of assignment.

An open question about this research is how well the findings of this paper can translate to other classroom settings. For instance, courses about technical subjects such as programming or math contain more abstract concepts than the course about sustainability used in this study; it is unclear how well the metrics introduced in this paper could aid instructors in predicting the understanding of learners in such courses. It will be important to collect data across various domains and levels of education in order to properly address concerns about the scalability of such metrics beyond the single course used in this analysis.

Our findings indicate that the novel metrics introduced in this paper have potential as predictive indicators of learner understanding of course material. Such information could be useful to instructors in a variety of different ways. For instance, instructors could use the results of these metrics in order to identify potentially struggling learners without having to assess each learner's concept map qualitatively. Additionally, characterizing what learners know at a given moment in time could help instructors to assemble project or study groups based on overlaps (or non-overlaps) in knowledge or interest.

There is also potential to use the collective network of merged learner concept maps as a pedagogical tool. For example, learners could research key nodes and/or edges that they overlooked but that other learners included, and write up a description of the nature of the missed content. Another possibility is to allow learners to browse an online visualization of the collective network to promote open-ended reflection. This experience allows learners to reflect on how their individual conceptual understanding connects to those of their peers. Having learners actively participate in the process of network construction, with each learner or group of learners taking responsibility for one piece of the final network, is in line with past CSCW research that emphasizes the human role in knowledge management [1], and can also be viewed as a form of Collaborative Sensemaking [67]. While the concept mapping assignments themselves are performed individually, by collecting mental models from learners in a standardized format, these assignments enable downstream collaboration and interplay with existing resources [26].

Finally, the current work does not address potential issues introduced by varying levels of concept specificity, a generally challenging task in concept mapping. Future steps to address this issue may take the form of enhancing network visualizations to emphasize chains of intermediary nodes

between two broad concepts, or using external resources such as Wikipedia to suggest potential alternatives for concepts that are too broad to be informative in a given context.

6 CONCLUSION

Network Analysis is shown to be a promising technique for analyzing the concepts and relationships in collaboratively-constructed educational networks. Typically, these networks have been composed of learners and their interactions, and it has been shown that learners occupying central positions within these networks tend to have better learning outcomes. This paper answers the call for analyses of more diverse types of networks within the CSCL space, and at the same time contributes to the relatively unexplored area of Network Analysis over Novakian concept map data. In particular, this paper introduces the novel methodology of merging learner concept maps into a collective network and using this collective network to calculate centrality and consensus-based metrics for individual learners. When evaluated on two academic quarters of concept mapping data, these novel metrics are shown to be more significantly correlated with learner performance and more reproducible across datasets than the metrics traditionally used to evaluate concept maps in the literature.

While it is clear that network metrics can, in some sense, predict learner understanding of the material, it is also clear that the complexity of concept map-based networks will require more work to understand which metrics (or combinations of metrics) are the best predictors of learner conceptual understanding in a given context. Moving forward, it will be important to perform such studies across larger and more diverse datasets, in order to further evaluate the suitability of these metrics, and Network Analysis as a whole, for predicting learner understanding.

7 ACKNOWLEDGMENTS

This work was supported by the National Science Foundation award number 2121572.

REFERENCES

- [1] Mark S Ackerman, Juri Dachtera, Volkmar Pipek, and Volker Wulf. 2013. Sharing knowledge and expertise: The CSCW view of knowledge management. *Computer Supported Cooperative Work (CSCW)* 22, 4 (2013), 531–573.
- [2] Gulgun Afacan Adanir. 2019. Detecting topics of chat discussions in a computer supported collaborative learning (CSCL) environment. *Turkish Online Journal of Distance Education* 20, 1 (2019), 96–114.
- [3] Nitin Agarwal and Faysal Ahmed. 2017. Developing collective learning extension for rapidly evolving information system courses. *Education and Information Technologies* 22, 1 (2017), 7–37.
- [4] Payam Aminpour, Steven A Gray, Antonie Jetter, and Philippe J Giabbanelli. 2018. Is the crowd wise enough to capture systems complexities? An exploration of wisdom of crowds using Fuzzy Cognitive Maps. (2018).
- [5] Mary Besterfield-Sacre, Jessica Gerchak, Mary Rose Lyons, Larry J Shuman, and Harvey Wolfe. 2004. Scoring concept maps: An integrated rubric for assessing engineering education. *Journal of Engineering Education* 93, 2 (2004), 105–115.
- [6] Gautam Biswas, Krittaya Leelawong, Daniel Schwartz, Nancy Vye, and The Teachable Agents Group at Vanderbilt. 2005. Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence* 19, 3-4 (2005), 363–392.
- [7] Stephen P Borgatti, Ajay Mehra, Daniel J Brass, and Giuseppe Labianca. 2009. Network analysis in the social sciences. *Science* 323, 5916 (2009), 892–895.
- [8] Rita Cadima, Jordi Ojeda Rodríguez, and José María Monguet Fierro. 2012. Social networks and performance in distributed learning communities. *Educational technology and society* 15, 4 (2012), 296–304.
- [9] AJ Cañas, JD Novak, and FM González. 2004. Varieties of concept mapping. (2004).
- [10] Alberto J Cañas, Greg Hill, Roger Carff, Niranjani Suri, James Lott, Gloria Gómez, Thomas C Eskridge, Mario Arroyo, and Rodrigo Carvajal. 2004. CmapTools: A knowledge modeling and sharing environment. (2004).
- [11] Karina L Cela, Miguel Ángel Sicilia, and Salvador Sánchez. 2015. Social network analysis in e-learning environments: A preliminary systematic review. *Educational Psychology Review* 27, 1 (2015), 219–246.
- [12] Li-Chen Cheng and Hui-Chun Chu. 2019. An innovative consensus map-embedded collaborative learning system for ER diagram learning: Sequential analysis of students' learning achievements. *Interactive Learning Environments* 27, 3 (2019), 410–425.

- [13] Hichang Cho, Geri Gay, Barry Davidson, and Anthony Ingrassia. 2007. Social networks, communication styles, and learning performance in a CSCL community. *Computers & education* 49, 2 (2007), 309–329.
- [14] Laura H Clayton. 2006. Concept mapping: an effective, active teaching-learning method. *Nursing education perspectives* 27, 4 (2006), 197–203.
- [15] Simone CO Conceição and Linda D Taylor. 2007. Using a constructivist approach with online concept maps: Relationship between theory and nursing education. *Nursing education perspectives* 28, 5 (2007), 268–275.
- [16] Kalani Craig, Joshua Danish, Megan Humburg, Cindy Hmelo-Silver, Maksymilian Szostalo, and Ann McCranie. 2021. Net. Create: Network visualization to support collaborative historical knowledge building. *International Journal of Computer-Supported Collaborative Learning* 16, 2 (2021), 185–223.
- [17] Jingjing Cui and Shengquan Yu. 2019. Fostering deeper learning in a flipped classroom: Effects of knowledge graphs versus concept maps. *British Journal of Educational Technology* 50, 5 (2019), 2308–2328.
- [18] Marielle Dado and Daniel Bodemer. 2017. A review of methodological applications of social network analysis in computer-supported collaborative learning. *Educational Research Review* 22 (2017), 159–180.
- [19] Harry S Delugach, Letha H Etzkorn, Sandra Carpenter, and Dawn Utley. 2016. A knowledge capture approach for directly acquiring team mental models. *International Journal of Human-Computer Studies* 96 (2016), 12–21.
- [20] Jana Dlouhá and Simon Burandt. 2015. Design and evaluation of learning processes in an international sustainability oriented study programme. In search of a new educational quality and assessment method. *Journal of Cleaner Production* 106 (2015), 247–258.
- [21] Katherine Dunsmore, Jennifer Turns, and Jessica M Yellin. 2011. Looking toward the real world: Student conceptions of engineering. *Journal of Engineering Education* 100, 2 (2011), 329–348.
- [22] Tanja Engelmann and Friedrich W Hesse. 2010. How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning* 5, 3 (2010), 299–319.
- [23] Ernesto Estrada and Naomichi Hatano. 2008. Communicability in complex networks. *Physical Review E* 77, 3 (2008), 036111.
- [24] Mohammadreza Farrokhnia, Héctor J Pijera-Díaz, Omid Noroozi, and Javad Hatami. 2019. Computer-supported collaborative concept mapping: The effects of different instructional designs on conceptual understanding and knowledge co-construction. *Computers & Education* 142 (2019), 103640.
- [25] Albert Fornells, Zaida Rodrigo, Xari Rovira, Mónica Sánchez, Ricard Santomà, Francesc Teixidó-Navarro, and Elisabet Golobardes. 2015. Promoting consensus in the concept mapping methodology: an application in the hospitality sector. *Pattern Recognition Letters* 67 (2015), 39–48.
- [26] Hayden Freedman, André van der Hoek, and Bill Tomlinson. 2022. Improving Wikidata with Student-Generated Concept Maps. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 205–215.
- [27] Hong Gao, E Shen, Susan Losh, and Jeannine Turner. 2007. A review of studies on collaborative concept mapping: What have we learned about the technique and what is next? *Journal of Interactive Learning Research* 18, 4 (2007), 479–492.
- [28] Hannie Gijlers, Nadira Saab, Wouter R Van Joolingen, Ton De Jong, and Bernadette HAM Van Hout-Wolters. 2009. Interaction between tool and talk: how instruction and tools support consensus building in collaborative inquiry-learning environments. *Journal of Computer Assisted Learning* 25, 3 (2009), 252–267.
- [29] Alyssa W Goldman and Mary Kane. 2014. Concept mapping and network analysis: An analytic approach to measure ties among constructs. *Evaluation and Program Planning* 47 (2014), 9–17.
- [30] Wenge Guo and Joseph Romano. 2007. A generalized Sidak-Holm procedure and control of generalized error rates under independence. *Statistical applications in genetics and molecular biology* 6, 1 (2007).
- [31] Aric Hagberg, Pieter Swart, and Daniel S Chult. 2008. *Exploring network structure, dynamics, and function using NetworkX*. Technical Report. Los Alamos National Lab.(LANL), Los Alamos, NM (United States).
- [32] Caroline A Haythornthwaite. 1999. Networks of information sharing among computer-supported distance learners. (1999).
- [33] Cindy E Hmelo-Silver and Heisawn Jeong. 2021. An overview of CSCL methods. *International Handbook of Computer-Supported Collaborative Learning* (2021), 65–83.
- [34] Kathryn W Jablokow, Joanna F DeFranco, and Sally Sue Richmond. 2013. A statistical study of concept mapping metrics. In *2013 ASEE Annual Conference & Exposition*. 23–105.
- [35] Jeonghyun Kim, Hyeyun Lee, Yesom Yoo, Hanall Sung, Il-Hyun Jo, and Yeonjeong Park. 2015. Towards Smart Asynchronous Discussion Activity: Using Social Network Analysis to Investigate Students' Discussion Patterns. In *Emerging issues in smart learning*. Springer, 361–370.
- [36] Ian M Kinchin, David B Hay, and Alan Adams. 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 (2000), 43–57.

- [37] Vassilis Komis, Nikolaos Avouris, and Christos Fidas. 2002. Computer-supported collaborative concept mapping: Study of synchronous peer interaction. *Education and Information Technologies* 7 (2002), 169–188.
- [38] Ismo T Koponen and Maija Nousiainen. 2019. Pre-service teachers' knowledge of relational structure of physics concepts: finding key concepts of electricity and magnetism. *Education Sciences* 9, 1 (2019), 18.
- [39] Maarten De Laat, Vic Lally, Lasse Lipponen, and Robert-Jan Simons. 2007. Online teaching in networked learning communities: A multi-method approach to studying the role of the teacher. *Instructional Science* 35, 3 (2007), 257–286.
- [40] Peter Lampert, Martin Scheuch, Peter Pany, Bernhard Müllner, and Michael Kiehn. 2019. Understanding students' conceptions of plant reproduction to better teach plant biology in schools. *Plants, People, Planet* 1, 3 (2019), 248–260.
- [41] Leigh Anne Liu, Ray Friedman, Bruce Barry, Michele J Gelfand, and Zhi-Xue Zhang. 2012. The dynamics of consensus building in intracultural and intercultural negotiations. *Administrative Science Quarterly* 57, 2 (2012), 269–304.
- [42] Sannyuya Liu, Lingyun Kang, Zhi Liu, Jing Fang, Zongkai Yang, Jianwen Sun, Meiyi Wang, and Mengwei Hu. 2021. Computer-supported collaborative concept mapping: the impact of students' perceptions of collaboration on their knowledge understanding and behavioral patterns. *Interactive Learning Environments* (2021), 1–20.
- [43] Kimberly Maguire Markham. 1992. *The structure and use of biological knowledge about mammals in novice and experienced students*. Ph.D. Dissertation. University of North Carolina at Wilmington.
- [44] Daniel McLinden. 2013. Concept maps as network data: Analysis of a concept map using the methods of social network analysis. *Evaluation and program planning* 36, 1 (2013), 40–48.
- [45] Brian Moon, C Johnston, and S Moon. 2018. A case for the superiority of concept mapping-based assessments for assessing mental models. In *Concept Mapping: Renewing Learning and Thinking. Proceedings of the 8th Int. Conference on Concept Mapping, Medellín, Colombia: Universidad EAFIT*.
- [46] Joseph D Novak. 1990. Concept mapping: A useful tool for science education. *Journal of research in science teaching* 27, 10 (1990), 937–949.
- [47] Kari A Nurmela, Erno Lehtinen, and Tuire Palonen. 1999. Evaluating CSCL log files by social network analysis. (1999).
- [48] Debra L O'Connor, Tristan E Johnson, and Mohammed K Khalil. 2004. Measuring team cognition: Concept mapping elicitation as a means of constructing team shared mental models in an applied setting. (2004).
- [49] Fan Ouyang. 2021. Using three social network analysis approaches to understand computer-supported collaborative learning. *Journal of Educational Computing Research* 59, 7 (2021), 1401–1424.
- [50] Claudia Picardi, Annamaria Goy, Daniele Gunetti, Giovanna Petrone, Marco Roberti, Nuninger Walter, et al. 2020. I learn. You learn. We learn? An experiment in collaborative concept mapping.. In *CSEDU 2020-12th International Conference on Computer Supported Education*, Vol. 2. Scitepress, 15–25.
- [51] Goran Putnik, Eric Costa, Cátia Alves, Hélio Castro, Leonilde Varela, and Vaibhav Shah. 2016. Analysing the correlation between social network analysis measures and performance of students in social network-based engineering education. *International Journal of Technology and Design Education* 26, 3 (2016), 413–437.
- [52] Debbie Denise Reese. 2004. Assessment and concept map structure: The interaction between subscores and well-formed mental models. In *meeting of the American Educational Research Association, San Diego*.
- [53] Pablo Reyes and Pierre Tchounikine. 2017. Mining learning groups' activities in Forum-type tools. In *Computer Supported Collaborative Learning 2005: The Next 10 Years!* Routledge, 509–513.
- [54] Scott R Rosas. 2017. Group concept mapping methodology: toward an epistemology of group conceptualization, complexity, and emergence. *Quality & quantity* 51, 3 (2017), 1403–1416.
- [55] Mohammed Saqr and Sonsoles López-Pernas. 2022. How CSCL roles emerge, persist, transition, and evolve over time: A four-year longitudinal study. *Computers & Education* 189 (2022), 104581.
- [56] Mohammed Saqr, Olga Viberg, and Henriikka Vartiainen. 2020. Capturing the participation and social dimensions of computer-supported collaborative learning through social network analysis: which method and measures matter? *International Journal of Computer-Supported Collaborative Learning* 15, 2 (2020), 227–248.
- [57] Beat A Schwendimann. 2014. Making sense of knowledge integration maps. *Digital knowledge maps in education: Technology-enhanced support for teachers and learners* (2014), 17–40.
- [58] David Williamson Shaffer, David Hatfield, Gina Navoa Svarovsky, Padraig Nash, Aran Nulty, Elizabeth Bagley, Ken Frank, André A Rupp, and Robert Mislevy. 2009. Epistemic network analysis: A prototype for 21st-century assessment of learning. *International Journal of Learning and Media* 1, 2 (2009).
- [59] Cynthia SQ Siew. 2019. Using network science to analyze concept maps of psychology undergraduates. *Applied Cognitive Psychology* 33, 4 (2019), 662–668.
- [60] Cynthia SQ Siew. 2020. Applications of network science to education research: Quantifying knowledge and the development of expertise through network analysis. *Education Sciences* 10, 4 (2020), 101.
- [61] Cynthia SQ Siew, Dirk U Wulff, Nicole M Beckage, and Yoed N Kenett. 2019. Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity* 2019 (2019).
- [62] Tuyi Siqin, Jan Van Aalst, and Samuel Kai Wah Chu. 2015. Fixed group and opportunistic collaboration in a CSCL environment. *International Journal of Computer-Supported Collaborative Learning* 10, 2 (2015), 161–181.

- [63] Gerry Stahl. 2016. From intersubjectivity to group cognition. *Computer Supported Cooperative Work (CSCW)* 25, 4 (2016), 355–384.
- [64] Neli Stoyanova and Piet Kommers. 2001. Learning effectiveness of concept mapping in a computer supported collaborative problem solving design. In *First European International Conference on Computer-Supported Collaborative Learning*, Vol. 13.
- [65] Binod Sundararajan. 2010. Emergence of the most knowledgeable other (mko): Social network analysis of chat and bulletin board conversations in a CSCL system. *Electronic Journal of e-Learning* 8, 2 (2010), 191–208.
- [66] Stefan Trausan-Matu and James D Slotta. 2021. Artifact analysis. In *International Handbook of Computer-Supported Collaborative Learning*. Springer, 551–567.
- [67] Karthikeyan Umapathy. 2010. Requirements to support collaborative sensemaking. In *CSCW CIS Workshop*, Vol. 10.
- [68] Sigrun Vehling, Carmine Malfitano, Joanna Shnall, Sarah Watt, Tania Panday, Aubrey Chiu, Anne Rydall, Camilla Zimmermann, Sarah Hales, Gary Rodin, et al. 2017. A concept map of death-related anxieties in patients with advanced cancer. *BMJ supportive & palliative care* 7, 4 (2017), 427–434.
- [69] Nicola Walshe. 2008. Understanding students' conceptions of sustainability. *Environmental Education Research* 14, 5 (2008), 537–558.
- [70] Duncan J Watts and Steven H Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature* 393, 6684 (1998), 440–442.
- [71] Zihan Wei, Lei Ren, Chao Liu, Mi Cao, Qun Yang, and Yanchun Deng. 2020. The concept map of felt stigma in patient with epilepsy. *Seizure* 80 (2020), 138–142.
- [72] Yue Yin, Jim Vanides, Maria Araceli Ruiz-Primo, Carlos C Ayala, and Richard J Shavelson. 2005. Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching* 42, 2 (2005), 166–184.
- [73] Qian Zhang, Qingtang Liu, Ni Zhang, and Linjing Wu. 2019. Social Network Analysis of Knowledge Construction in Computer-Supported Collaborative Learning. 399–402. <https://doi.org/10.35745/ecei2019v2.102>
- [74] Si Zhang, Qianqian Gao, Mengyu Sun, Zhihui Cai, Honghui Li, Yanling Tang, and Qingtang Liu. 2022. Understanding student teachers' collaborative problem solving: Insights from an epistemic network analysis (ENA). *Computers & Education* 183 (2022), 104485.

Received January 2023; revised July 2023; accepted November 2023