Refining Concept Maps as Method to Assess Learning Outcomes Among Engineering Students

Dr. Sean Michael Ferguson, University of Virginia

Sean Ferguson is a Lecturer in the Department of Engineering and Society at UVA since 2014. He specializes in sustainable technology, standards, and bioeconomy from a background in Science and Technology Studies from Rensselaer Polytechnic Institute, working on energy and environmental policy in New York State, and a former life in cellular biology.

Dr. Rider W. Foley, University of Virginia

Dr. Rider W. Foley is an assistant professor in the science, technology & society program in the Department of Engineering and Society at the University of Virginia. He is the principal investigator at University of Virginia on the '4C Project' on Cultivating Cultures of Ethical STEM education with colleagues from Notre Dame, Xavier University and St. Mary's College. He is also the co-leader of the 'Nano and the City' thematic research cluster for the Center for Nanotechnology in Society at Arizona State University. Rider is a Research Collaborator with the Sustainability Science Education program at the Biodesign Institute. His research focuses on wicked problems that arise at the intersection of society and technology. Rider holds a Ph.D. in Sustainability from Arizona State University, and a Master's degree in Environmental Management from Harvard University and a Bachelor's degree in Environmental Science from University of New Hampshire. Before earning his doctorate, he has worked for a decade in consulting and emergency response for Triumvirate Environmental Inc.

Mr. John Kofi Eshirow Jr.
Miss Catherine Claire Pollack, University of Virginia

Catherine Pollack is a fourth year undergraduate Rodman Scholar at the University of Virginia pursuing a double major in biomedical engineering and applied statistics. Following graduation, she will attend the Geisel School of Medicine at Dartmouth College to pursue a Ph.D. in quantitative biomedical sciences beginning Fall 2018.

Refining Concept Maps as a Method to Assess Learning Outcomes among Engineering Students

1.0 Introduction

Concept mapping activities have been used extensively in over 500 educational research projects with the goal of developing curricula, assessment, and testing knowledge acquisition [1]. This suite of methods, many shared among ASEE members, are proven to perform in a variety of settings and learning communities. Concept maps (CMs) are most often used to link course learning goals to individual student's knowledge integration of course material, especially where there are defined concepts and linkages between concepts that should be replicable by the students. By design, CMs also offer flexibility for students to explore the relationships between course and lived experiences, offering assessment opportunities to determine what students bring into a course and what they take with them as they progress through the curriculum. Further, concept mapping is widely accepted as encouraging improved learning experiences in comparison to, or in conjunction with, traditional teaching methods [2].

However, concept maps are less well understood as an approach to understand knowledge acquisition and competency for representing complex and dynamic interactions between socio-cultural and technological systems. . a static body of knowledge from a textbook. Learning assessments are confounded by such ambiguity, which opens up questions about how to determine what amounts to a "good" concept map. This is particularly evident when student-generated concept maps cannot be analyzed against an absolute target,. Further, without the ability to define hierarchies of key concept to sub-concept in dynamic socio-technical systems, there is a challenge to assess the orientation of knowledge acquisition for students [3], [4]. This research considers traditional scoring of concept maps that tend to emphasize node and connection quantity [5] (i.e., the number of concepts expressed), which might be problematic for liberal arts courses demanding engineering students critically reflect and rethink their prior assumptions and heuristics about the relationship between technology and society. There is a need to reliably capture student learning about complex and dynamic socio-technical systems without privileging an assessment tool that *a priori* evaluates "more is better".

With that in mind this manuscript addresses three key issues in this area. The first objective, *efficiency*, is to interrogate the use of concept maps to capture student learning about the complexity of socio-technical systems in large-scale engineering programs where a review of each individual map would require extensive time investments. Conducting the concept mapping exercise and analysis strategy are impacted under this objective. This leads to the second objective, *methodological development,* which assesses how complexity can be evaluated in the quantity and structure of the concept maps, regardless of substantive content, see [5] for previous studies on content development. This two-fold objective prompts our first question: Can aggregate datasets reveal meaningful changes in student learning without accounting for substance and content?

The third objective, *a prior knowledge*, explores how a student's prior knowledge and changes overtime (beyond one course) is an important factor when evaluating an engineer's

perceptions of technology-society relationships. As we explore this issue, we call into question the assumption that students are a "blank slate" coming into a course. To this end we ask: Can concept maps document prior knowledge and then show changes in how representations of knowledge change after a course? Might knowledge displacement be a worthwhile outcome, if the student's initial understanding is fraught with unfounded heuristics? Since liberal arts courses for engineers draw from many domains beyond the technical (e.g., politics and cultural anthropology), we explore questions of whether instruction is displacing previous acquired knowledge, refining current understandings, or supporting a net increase in knowledge acquisition.

This paper aims to address those main points, while transparently sharing our experiences and strategies in performing concept mapping in an efficient and rigorous manner. This paper offers both a methodological evaluation of student learning about socio-technical systems that is part of an unfolding, iterative research project with the ultimate goal of using CMs as an assessment tools for longitudinal learning outcomes. Linking to previous publications assessing professional skills[1] in engineering education, this research paper provides a comparison of the methods employed in academic years 2015-16 and 2016-17 from multiple classes within a single department to assess student outcomes through the use of concept mapping activities. The department, Engineering and Society at the University of Virginia, is interdisciplinary and is considering new assessment tools to match ABET accreditation objectives [6]. This paper also shares alternative forms of data analysis and reveals how minor changes affect the validity of CMs as a platform for internal assessment of students' knowledge acquisition as compared to prior literature. What follows is a systematic review of prior literature on knowledge acquisition, displacement, and retention, as well as concept mapping methodology. We then review the research design, course context and forms of data gathering and analysis.

2.0 Literature review

Concept maps have been repeatedly used in education due to their usefulness for learning and assessment. Concept maps can we designed for free association, documenting complexity, and contextualize when ideas emerge and travel, and incorporate lived experiences [8]. Documenting a set of concepts is one facet of the concept mapping process. Shifts in the relationships between concepts due to experiences in and outside of the class and the freedom to make connections between subject-areas creates an opportunity for understanding meta-cognitive learning. Analyzing concept maps by students over a time period can yield metrics on higher complexity scores, more extensive hierarchies, and appreciation of concept linkages and feedback loops [5]. For example, previous studies that compare post-course maps to pre-course maps show that students become knowledgeable about subjects they had little in the way of experience with before attending the course [7]. Course specific testing—e.g. topically

---

[1] For a better understanding of the use of "professional skills" rather than "soft skills" please see K. Neeley's ASEE Conference proceeding on the gendered and derogatorily characteristics of the "soft skills" phrase in a male dominated field.

constrained quizzes—do not offer this same level of opportunity to understand the intellectual history and emotional disposition of students as they engage in new learning activities [9].

As Kinchin [8] notes in a meta-analysis of several decades worth of publications on concept maps, flexibility in adapting CMs to new scenarios and different platforms can be achieved without degrading the reliability of the basic mechanism of capturing knowledge acquisition and documenting learning outcomes. Important for our analysis, Gouveia and Valadares [9] document how concept maps not only indicate learning due to a given course experience, but also provide insight about what students bring into a course. The authors argue, "Concept mapping is a technique that exposes the concepts and assertions hidden within the cognitive structure of each student and it is of great importance since it shows the changes occurring in this cognitive structure, clarifies misconceptions and superficial interpretation in the teaching/learning process, and allows the teacher and student to exchange points of view on the validity or absence of a link between two concepts [10, p2]." Thus, concept maps can give instructors information about how to best scaffold the learning experience for the students in a particular classroom and to conduct an initial assessment of what is already known and what gaps are present at the onset of the course.

The following section reviews prior studies shared at ASEE documented knowledge acquisition in technical and professional skills in engineering curriculum [5]. This provides background on the strategies deployed to understand knowledge acquisition in epistemologically complex courses and the importance of documenting prior knowledge and the procedures by which knowledge might be displaced.

2.1   Knowledge acquisition, displacement, and retention

Several conceptual foundations guide the researchers understanding of key knowledge changes among students who experience similar course designs to STS4500/4600 and the assessments of learning. Shankar and colleagues took a case study approach to bring aspects of the humanities into engineering education [11]. The course goal was to expose engineering students to social and community issues. They believed that having a course that tackled issues faced in underrepresented communities would help them recruit more women and minorities. The pedagogical theory employed allowed the students to learn more because they would be interacting more with each other while hearing about real-world issues that are out of their comfort zones, such as police homicide rates on minorities in metropolitan cities. This case illustrates that engineers can acquire new knowledgeable about the broader societal context if they engage in interdisciplinary liberal arts education. The Science, Technology, and Society (STS) courses at the UVA share traits with this style of curriculum. Importantly for this study, pre- and post-surveys show an increase in technical and professional skills that align well with the ABET EC 2000 Criteria under evaluation.

David and Marshall [12] claim that students acquire knowledge when they have opportunities to explore open-ended topics where students select topics they deem important. David and Marshall employed project-based learning techniques that created near-term milestones that culminated in final project presentations. Students set goals, assessed their

progress and reflected upon their performance. This provided more autonomy and participation, thus motivating students to be more active in their own learning. Students demonstrated knowledge acquisition by applying lessons from the course in their 'real-world' projects. Students were able to learn from their mistakes, and from their attempts to problem solve. The course offered clear metrics of success in both technical and professional skills designated by the ABET EC 2000 criteria. Yet, the authors critically reflected on the lack of clear assessment measures for engineering fundamentals and the challenge of scaling this course to meet the demands of larger enrollments.

While an efficient method for quickly distributing information to students, traditional lecturing has been observed to be an ineffective method for engaging in subjective and dynamic topics. Student engagement quickly wanes, as they tend to fall asleep or become confused in how to situate the information into real world experiences. Garlock and colleagues sought to teach literacy through an interactive course design that entailed lecture-based demonstrations, hands-on activities, polling questions, and discussion questions [13]. This course offered three phases of interaction: predict, experience, and reflect. The classes started with an online poll that gave students an opportunity to share their preconceived notions, heuristics, and knowledge about the topic of the day. This provided the instructor with an understanding of the students' prior knowledge. The "experiencing" stage involved an interactive lecture or in-class demonstrations. This allowed students to see, touch, or otherwise discover new knowledge during the 'experience'. In the final stage students reflected on how the concept manifests in the 'real world'. This paper highlights the notion that, while students enter the classroom with preconceived notions about a given topic, engagement with these pre-conceptions assist the instructor in facilitating the displacement or reformulation of misguided notions, augmentation of existing knowledge, and the development of connections between the topic and the students' lived experience in the 'real world'. However, assessing this process is resource intensive.

2.2   Methodological Approaches to Concept Maps

CMs are designed to graphically document a student's knowledge of smaller concepts and make linkages to other concepts and thus represent a broader, meta-concept. Each concept is contained within a single node and the links (or connections) indicate relationships, which are typically labeled with verbs or other explanatory text [14]. Application of this technique can take many forms in documenting knowledge or changing understanding of a subject. CMs are typically integrated into teaching as either learning tools or documentation of conceptual understanding over a period of time. Resulting CMs depict growth in student understanding or offer opportunities for an instructor to step in to correct confusion. Alternatively, CMs can inform curriculum design and assessment strategies. We will not go into an in-depth analysis of the merits of concept mapping as this can be found in our previous publication and meta-analysis of concept mapping use, such as Nesbit and Adenosope's [2] and Kinchin's [7] evaluation of several decades of CM-related publications.

Several studies have compared student-generated CMs to expert exemplars [15]. Others have tracked the ability of students to fill in correct concepts based upon an incomplete expert scaffold or use all available concepts from a course [16]. As documentation of final learning

goals, CMs can be used as an alternative assessment tool to more traditional final exams [17], [18]. In all cases, the benefit of concept mapping comes from the possibility of highlighting the growing expertise in a given field as a dynamic knowledge framework and not a static set of isolated ideas [19].

Important for this study are the scholars who have integrated CMs into courses related to sustainability, systems thinking, green engineering and other interdisciplinary studies, which do not have "perfect" definitions that lend themselves to direct comparisons between experts and students. Segalàs et al. [20] investigated the effectiveness of concept maps to document the success of sustainability courses in several iterations and across multiple nations. Their coding scheme informed our own coding methods due to their experience in evaluating courses that had no "correct" final exemplar. Similarly, Borrego et al. [21] and Lourdel et al. [22] found improvements in the comprehensiveness and correctness of students understanding of engineering and sustainability after taking a course. Both studies note an increase in the number of nodes present alongside shifting student portrayals of superficial acknowledgement of economics and environment to more rigorous integration of social and technical complexity.

A question among scholars experimenting with concept mapping is what constitutes a "good" map. Avid proponents and developers of mapping software, Cañas, Novak, and Reiska [3] call into question previous depictions of quality in mapping. They find a discontinuity in how researchers have quantified concept maps. In particular, they argue that too often high values of concepts (aka nodes) are assumed to represent complexity or quality. Rather, CMs must be understood in their epistemological context. At an individual level assessment, in a course with predefined core concepts, comparing an expert developed map to a student generated map would offer useful details on which students are lacking understanding of course material. Often the concept mapping activity requires mappers to select from a given set of concepts to construct the correct interrelationships between concepts. Alternatively, successful memorization of a defined set of concepts and hierarchical structuring of concept to sub-concept becomes the most important evaluation of quality. Whereas Cañas et al. privilege this latter type of assessment, there are possibilities to explore deeper understanding of a smaller numbers of concepts by considering non-hierarchical cross-linking of difficult to understand ideas. Feedback loops and overlapping bridges between knowledge domains might be the most essential characteristic of a "good" map.

2.2.1 Prompts for concept maps

As Nesbit and Adenosope's [2] have noted, it is important to provide a central theme to start a concept mapping activity. There are "closed" prompts, in which a student is directed to re-create a correct, expert created concept that is core to the course. There are also more open-ended prompts that allow for the participant to construct a meta-concept around a central idea. Yin et al. [23] found that when comparing two approaches, the minimally directed prompts allowed the participants to more freely express their knowledge (or current understanding) of the subject. To this end, there are quite a few instances where the prompts afford participants to construct their own map around ideas such as "science policy" [24], "sustainability" [25], or "statics" [26]. For the most part, our review of the literature only revealed prompts that entailed

words or phrases. We note that there is room to explore whether participants shown an object or artifact might be worthwhile or could introduce variability as a prompt alternative.

2.2.2   Nodes, Connections, Density

One of the most basic approaches to evaluating concepts maps is to count the total number of nodes (smaller concepts) and connections (or linkages) between the nodes before calculating the levels of complexity [links (L)÷ nodes (N)], as shown by Jablokow and colleagues [27] in their 2013 ASEE paper on statistical methods for analyzing concept maps. The assumption is that by counting more nodes, the result indicates knowledge acquisition. This is where we want to problematize the 'more is better' assumption by evaluating the CMs from a set of interdisciplinary courses using several quantitative metrics. In an interdisciplinary course considering ethical scenarios or cultural relativity, where there is no clear indication of what is a "good" concept, simply having a greater number of nodes might not suggest greater knowledge integration or effective learning. As an additional consideration, based upon the literature reviewed above (section 2.1), we explore how the number of nodes might be an indicator of displacement or replacement of prior heuristics. Secondarily, we investigate how a result of more nodes may indicate the sharing of prior tacit knowledge learned outside of the course. The counting of nodes and connections does not address what was unlearned, supplanted, and what was gained. Thus, we question if knowledge acquisition cannot be understood by nodes, connections, and map density alone.

2.2.3   Concept Map: Hierarchy, form, and structure

Prior work on the hierarchy, structures, and forms of concept maps is quite robust. In regards to hierarchy, Kinchin et al. [28] employed three analytical levels: no hierarchy (spoke and wheel), multi-level hierarchy in a linear fashion (chain-link), and nets (or networks) that involve multi-level hierarchy and linkages among the branches. The work of Turns and colleagues [19] in engineering education interpreted the cross-linkages as feedback loops that bring dynamism to the concept maps. Thus, as McClure et al. [29] argued, the hierarchy and form of the concept maps is a secondary measure of map complexity that can be interpreted independently of quantification of nodes and connections as calculated by Jablokow et al [27].

3.0 Research Design and Methods

This research builds off prior work [5] that focused solely on the substance of the concept maps and developed a unique coding structure modified from Shallcross [25] to test student learning outcomes after a year-long course in their final year of studies. Evaluating pre- and post-test CMs demonstrated considerable shifts in the diversity of knowledge domains present and aligned those learning outcomes with ABET EC 2000 assessment criteria and the program's curriculum. Based upon that study and considering comments received on previous publications, this paper shares how we have refined our protocols and thus offer comparisons between new and published data sets. The protocol is evaluated alongside evaluation of the merits of the instrument for programmatic assessment. This research is then situated to re-evaluate a previous data set using a new method derived from previously published research evaluating CMs. Additionally, using this new method on a new data set from the same course as well as a

precursor course offers opportunities for refining a powerful educational assessment tool based upon what we have learned in this iterative process.

3.1 Concept Map Protocol and Course Context

The intent of our protocol was to avoid the time and resource expenditures involved in many software-based concept map programs that require significant time requirements and learning to give students the confidence and skills to perform the task at hand. Thus, we relied on lower-technology dependent application with the use of paper and pencil (or pen), which Muryanto [31] argues is an acceptable format and affords greater flexibility due to time and resource constraints. To that end, the procedure included a brief introduction to the mapping activity provided by a handout that also contained space on the reverse side of the page for the activity itself. A consent form and explanation were presented to the potential participants per our Institutional Review Board – Social and Behavioral Study Protocol. The protocol was implemented in one semester long course and one year-long sequence.

Science, Technology and Contemporary Issues (STS 1500) at UVA aims to introduce students to the relationship between engineering, technology, and society. The course provides students with an introduction to engineering ethics and the legal and social dimensions of engineering practice". STS 1500 is usually taken in the students' first year of engineering coursework at the university. The final two courses (STS 4500 and STS 4600) in the curriculum are taken in the students' final academic year of classes. The learning objectives for STS 4500 and STS 4600 (hereby referred to as STS 4500/4600) are, "engaging students with the challenge of framing and solving engineering problems in a manner that requires attention to social dimensions. Students are introduced to STS theories and methods as a means to prepare them for their STS research papers." The STS research papers are bound with the student's technical report and together constitute their senior thesis requirement for graduation. Given that student's technical research varies, that students utilize different STS frameworks and methods (given the interdisciplinary nature of the field). As many of the technologies they explore are emerging, there can be no "perfect" or "right" answer when assessing the relationship between technology and society. Each student's response will be contextualized by the three factors above (technical project, STS framing, and uncertain futures) and, likely, many life experiences internal and external to the course.

Thus, in an attempt to find common ground from which to start, participants (students in STS 1500 and STS 4500/4600) were shown one of two artifacts to start their concept maps: an iPod Touch (circa 2008) and a road bicycle. The use of a physical artifact is novel, as far as we are aware, and the justification for these two objects is justifiable given the intent of the mapping activity and previous student experiences. The iPod is a relatively small "black box" with very little outward facing complexity. It is small in size, is computer-oriented, and connects to the internet. This artifact has seen significant transformation within the students lived experience and is associated with a large and controversial company and figurehead. This offers a rich space to expand from as the participants consider the overall message to consider how to depict their thoughts about "science, technology, and society". The bike offers an open design, as the parts can be seen as mechanistically complex and components are readily identifiable. The bike is a

more mundane object and relates to different user types and markets. This artifact has experienced less change in the past decade several decades relative to the iPod, being made of standard steel construction and acting as the researcher's personal commuter bike. We constrained the time to work on the concept maps to fifteen minutes (15min), as observations during earlier pilot project showed that greater than 95% of students sat idling after that period of time when given twenty minutes.

The pre-test was conducted during the first three weeks of the course and the post-test was within the last two-weeks of the conclusion of the course. The change in timing on the post-test for STS4500/4600 was, in part, changed due to feedback from students in a focus group session where they stated they were 'checked out' on the last day of school in the spring 2016, since the concept map activity was not a graded assessment or exam (Table I). The post-test after spring break 2017 was ungraded but was called a 'summative assessment'; student participation was markedly different. Furthermore, the post-test in spring 2017 placed a stronger emphasis on the artifact in question, as non-compliance (not using the provided artifact in the concept map) was high in the spring 2016 post-test on the last day of their senior year. Samples were collected from three courses in the past two years for this study. The academic year, course titles, participants, student year (1 through 4), times of data collection are shown in Table I, below.

Table I. Sampled courses. Total participants in this study (n) 319.

Academic Year	Course Title	Participants	Level	Data Collection		
				Pre T0	Post T1	Post T2
2015-16	STS4500/4600	78	4th Year	1st Week	Last week Fall	Last week Spring
2016-17	STS 1500	152	1st Year	2nd week	Second to last day	n/a
2016-17	STS 4500/4600	89	4th Year	1st week	Post spring break	n/a

3.2 Data Analysis

3.2.1 Timing and differences in pre-post tests

The data analysis compares the differences in results by focusing initially on the simplistic measures of nodes as they relate to the timing of the pre- and post-tests. Differences in pre-post node counts (T0-T1) are assessed between students enrolled in a mandatory first-year course (STS 1500) and a second community of fourth-year student enrolled in a mandatory, year-long course (STS 4500/4600). Analysis of knowledge acquisition, as well as other statistical testing, was conducted and visualized with the R-statistical software package [32].

3.2.2 iPod versus Bike versus other (non-compliant maps)

All maps were coded for iPod, Bike and other (e.g. non-compliant objects) as categorical variables for this analysis. Boxplots for each category of pre-post-nodes were generated in order to better visualize the spread of the data and identify possible outliers (students who generated significantly more or less nodes compared to peers within their course). The boxplots also

display differences in mean scores for the pre- and post- maps based upon the object present for the students (Fig 4).

3.2.3 Hierarchies and structures

The measure of hierarchy and map structure was tested with two alternative scoring systems. The first scoring system, Hierarchy Structure A (Fig 1), references the system used by Foley and colleagues (2017) that employed a four-tier scale-score for the maps: 0 (no hierarchy), 1 (two levels of hierarchy), 2 (more than two levels of hierarchy), and 3 (greater than two levels of hierarchy and multiple feedback loops). This scoring system had been used successfully in documenting knowledge acquisition among preservice teachers in a concept mapping test. The second scoring system, Hierarchy Structure B, was developed as a methodological test for whether more nuance and diversity in the map structures might be hidden by too few categories of complexity, see **Error! Reference source not found.**Fig 2. below. Two undergraduates independently scored the concept maps and did not compare their results with each other before this analysis. The scoring systems entails:

1. Shows a linear structure with no branching and spoke and wheel up to two nodes from the center.
2. Captures three simple forms with up to 2 nodes from center: branching off central node, one interaction between branches, and interactions are on the same linear path.
3. Represents linear structures or spoke and wheel, extend beyond two nodes from center.
4. Builds upon the three simple forms, but the nodes exceed beyond two nodes from the center and contains a few (1 or 2) interactions across branches and can include feedback loops
5. Displays the most complex network with multiple branches from the center, greater than nodes beyond the center and multiple interactions and can contain feedback loops.
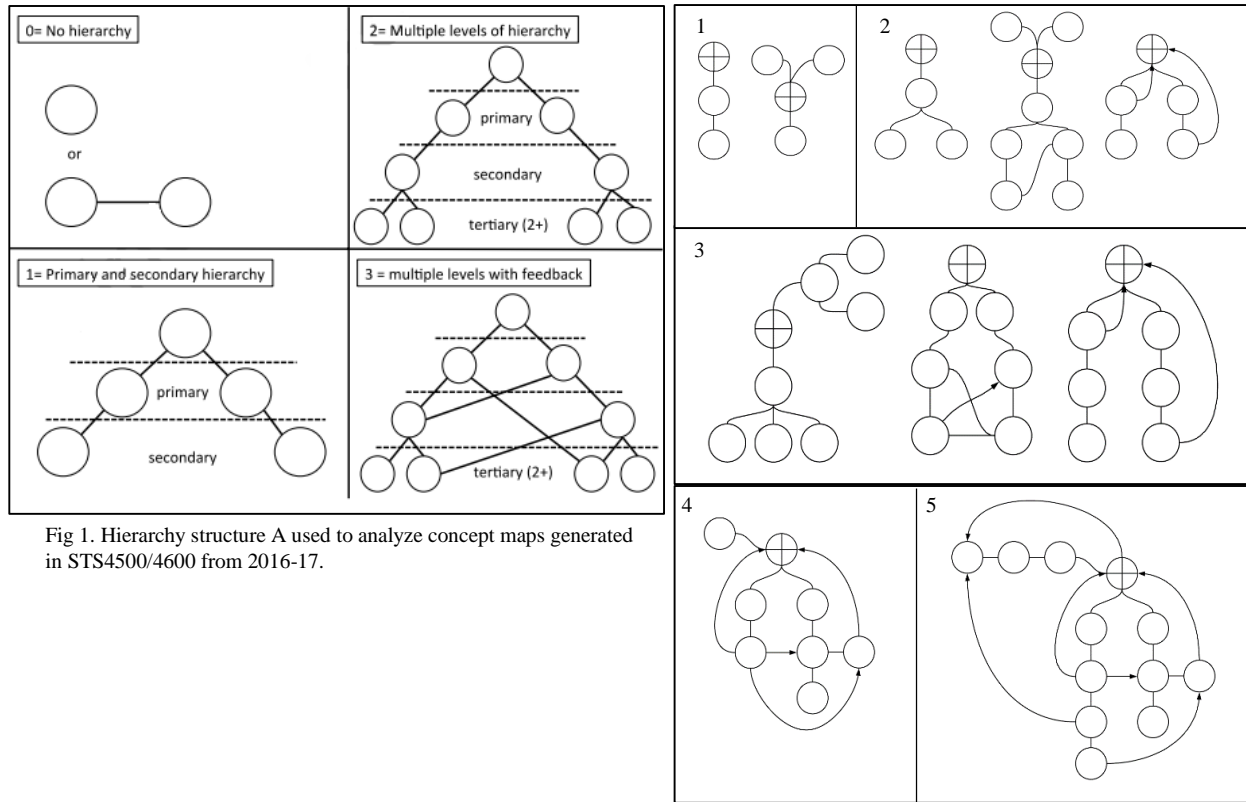
Fig 1. Hierarchy structure A used to analyze concept maps generated in STS4500/4600 from 2016-17.

Fig 2. Hierarchy structure B used to analyze concept maps generated in STS4500/4600 from 2015-16.

4.0   Results: Exploring methodological choices

CMs produced by first-year students in STS1500 shows little change between the pre- and post-tests, but with a slight negative trend (i.e., more pre-nodes than post-nodes). Yet, significant change is observed among the students that complete the pre-post tests in their fourth year in the STS4500/4600 course sequence (section 4.1). This opens up questions about how assumptions and heuristics that first-year students hold about technology-society relationships are addressed in the course and the receptivity of first-year students to this type of 'myth busting' content. The methodological questions regarding the technological artifact (bike or iPod) show no difference in pre-post nodes (section 4.2), while the method for analyzing the concept map structure clearly favors the more attenuated coding scheme shown in Fig 2 above, discussed in section 4.3.

4.1   Node count as key indicator of knowledge acquisition.

The three courses (STS 4500/4600 in 2015-2016, and STS 1500, STS 4500/4600 in 2016-2017) were plotted in a single scatterplot in order to compare pre-post node generation more easily across groups (Fig 3). The identity line was included on the graph as a way to visually compare how the amount of nodes changed for each student between the pre- and post-

CMs. Points that are located above the identity line indicate that students had more nodes on their post-CMs than their pre-CMs, while points located below the line suggest students had more pre-nodes than post-nodes. Solid lines are a representation of a smoothed, nonparametric trend line (LOESS: Locally Weighted Scatter-plot Smoothing), with each color line corresponding to the dots of the same category. The gray areas capture the 95% confidence intervals around each of the trend lines. The results of this test forces four questions to be addressed: whether the assessment heuristic that 'more is better' relates to knowledge acquisition since students generated post-test maps with less nodes depending on the course, epistemological disruption caused by courses intended to "shake up" student understanding of technology and society relationships, scaffolding from sequential course requirements supporting exploration of complexity, and which nodes should count as knowledge related to the course.

Upon closer inspection, the senior students (STS 4500/4600) account for a large majority of the increased counts (e.g., more post-nodes than pre-nodes, showcased by blue and green points), while the freshman students enrolled in STS 1500 represent the majority of those with decreased learning outcomes (e.g., more pre-nodes than post-nodes, shown in red points). This might suggest that the students had to reformulate their understanding of technology-society relationships. If this is the case, then a reduction in nodes might be indicative of knowledge disruption, but a productive disruption as the students are forced to confront assumptions, cognitive structures, and foreign ideas. An alternative explanation could be that students in STS 1500 were de-motivated by the course and therefore did not put forth a strong effort. Those questions will guide an upcoming focus group with outlier students that transitioned from many nodes to fewer and vice versa. The relative majority of positive changes between pre-post maps occurred among the seniors in STS 4500/4600. These students experienced at least two prior courses in the department as well as a diversity of courses, internships, and other extracurricular experiences. The ability to produce more nodes could be due to the previous scaffolding in evaluating engineering in context.
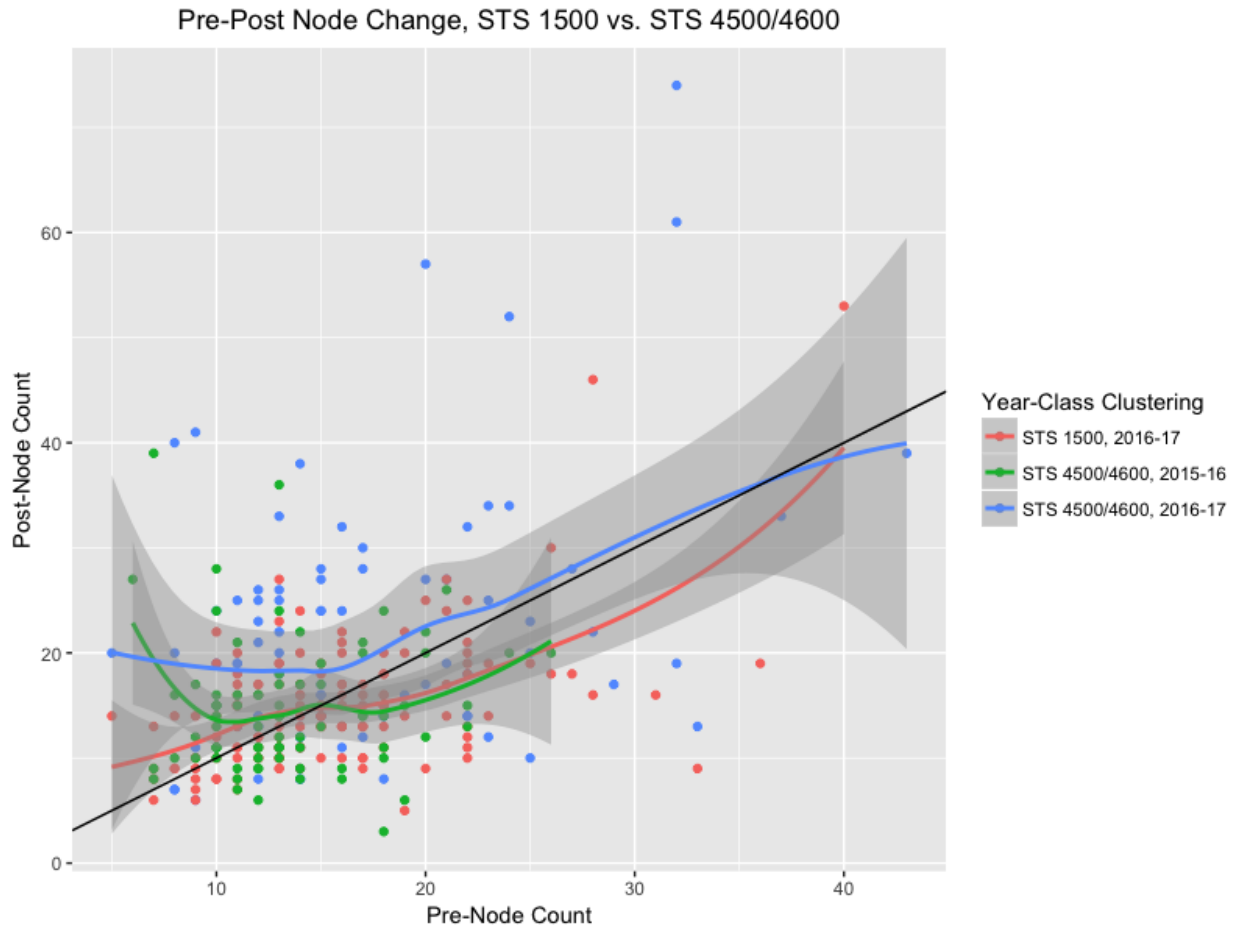
Fig 3. Pre-post node comparison for all students between 2015-2017 in STS 1500 and STS 4500/4600. The red dots represent first-year students in STS 1500 in 2016-17 and the red line reflects the nonparametric trend line. The green dots are data points from fourth-year students in STS4500/4600 in 2015-16, while the blue dots represent fourth-year students in STS4500/4600 enrolled in 2016-17. The green and blue lines are the nonparametric trend lines for those data points. The black line is the identity line, which in this case corresponds to no change between pre- and post-node count.

4.2   Bike or iPod: Does it matter?

In this study, we explored the use of objects, rather than words as the originating points for the concept maps. In order to ensure the artifacts themselves did not shift performance on the concepts maps a check on pre-post maps across all tested courses was completed. In the STS 1500 course there is no statistical difference in the pre-post nodes generated for each object (Fig 4). In the 2016-17 fourth-year course, the pre-test shows almost exactly the same number of nodes and there is no statistical difference in the post-test nodes between the objects (Fig 5). This indicates that student learning outcomes are captured evenly, regardless of the object used to initiate the concept mapping activity.
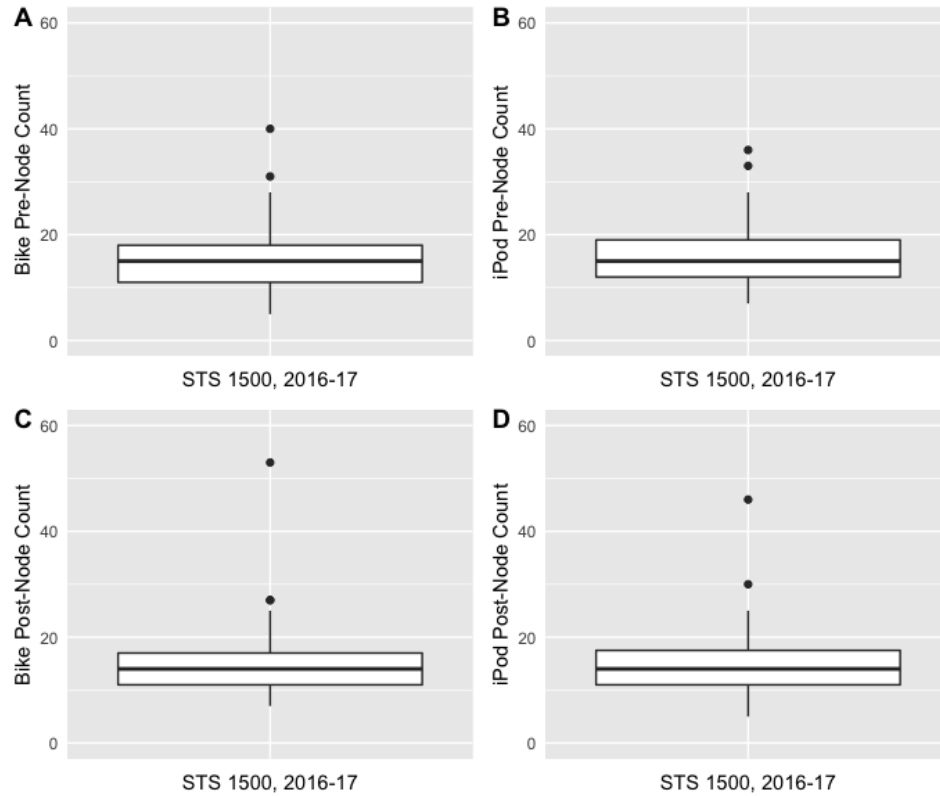
Fig 4. Box plots of pre-post concept maps using a bike or iPod as the starting node in STS1500. Rows indicate pre- or post-nodes, while columns indicate item. In particular, A represents bike pre-node counts (n = 94), B represents iPod pre-node counts (n = 55), C represents bike post-node counts (n = 94), and D represents iPod post-node counts (n = 55).
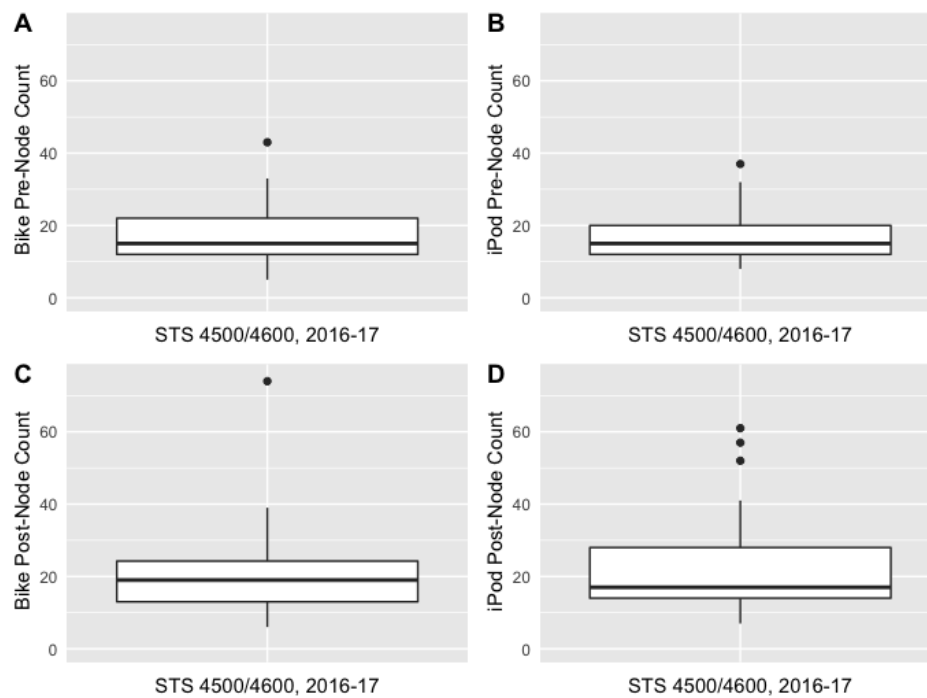

Fig 5. Box plots of pre-post concept maps using a bike or iPod as the starting node in STS 4500/4600 in the 2016-2017 school year. Rows indicate pre- or post-nodes, while columns indicate artifact. In particular, A represents bike pre-node counts (n = 44), B represents iPod pre-node counts (n = 45), C represents bike post-node counts (n = 44), and D represents iPod post-node counts (n = 45).

The 'other' maps were created when students did not comply with the instructions given and started their concept maps with random ideas, such as ice cream or basketball (during the NCAA tournament). While the non-compliant maps were originally dismissed due to the non-compliance, they offer interesting instances of artifact self-selection. The maps were included in analysis in case they demonstrated any unique characteristics. The mean pre-nodes generated for the bike is the lowest at 8.0, while the iPod pre-node mean is approximately 9.1, and the 'other' mean pre-node count is at 10.0 (Fig 6, below). Post-test means, which have a more balanced n, show a slightly different story where the iPod artifact node mean count (16.33) is largest, followed by the 'other' group (14.67) and the bike group (12.91). This "natural experiment" could indicate that when students choose their own object they express more creativity and thus generate greater nodes. Creativity has long been trumpeted as a learning outcome captured by concept maps and thus, the object placed in the center of the room certainly might constrain students' thinking and force them to start from a given point. While our initial determination is that the technical artifact might not matter in the aggregate for stimulating concepts, there is the possibility for determining if more open-ended mapping activities might encourage unforeseen ideas students find relevant to the curriculum, course, or their own experience as emerging engineers. Without further exploration, 'non-compliance' remains an indeterminate variable still requiring study.
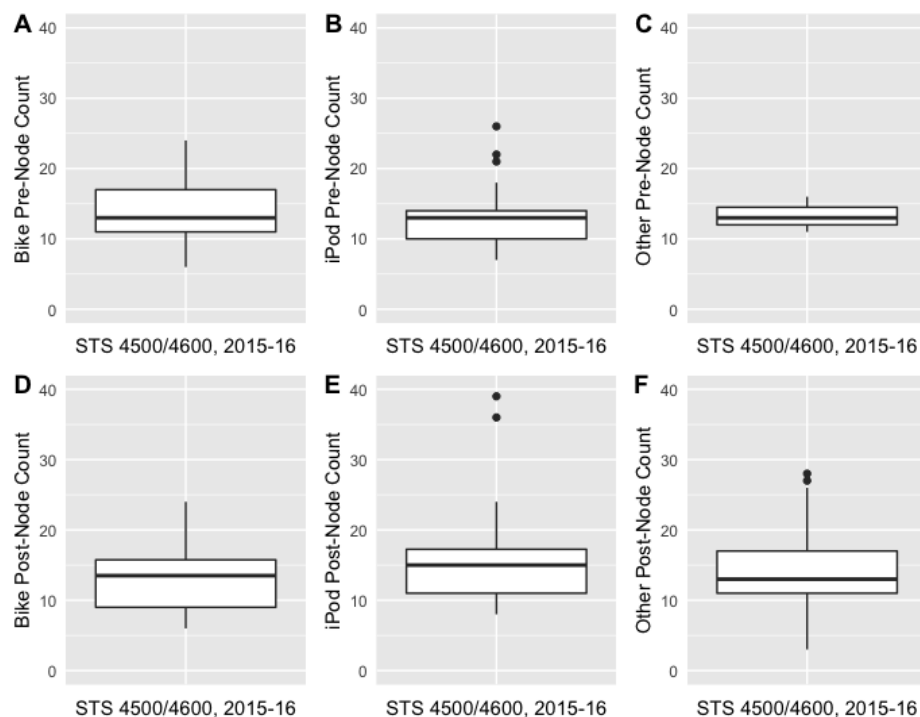

Fig 6. Comparison between bike, iPod, and non-compliant other. Data includes all concept maps generated from STS 4500/4600 between 2015-2016. Rows indicate pre- or post-nodes, while columns indicate artifact. In particular, A represents bike pre-node (n = 39), B represents iPod pre-node (n = 37), C represents non-compliant other pre-node (n = 3), D represents bike post-node (n = 22), E represents iPod post-node (n = 24), and F represents non-compliant other (n = 33).

4.3   Hierarchies and structure

Two separate measures of complexity were used to score the concept maps. The first of which was a four-level score 0-3 (Fig 1) that was applied to the 2016-17 concept maps generated in STS 4500/4600. That yielded a pre-test hierarchy score with a mean of 2.03 and a post-test hierarchy score of 2.09. A paired t-test for difference of means was conducted at a significance level of 0.05, yielding no significant difference between hierarchy structures. The second measure of hierarchy and structure entailed a five-level scoring system (Fig 2), which was applied to the 2015-16 concept maps generated in STS 4500/4600. The results for the second scoring system display a clear shift in the levels of hierarchy and complexity in structure (Fig 7, below). The line of best fit shows that the pre- maps (blue bars) trend toward less complexity, while the post-test maps (orange bars) trend toward greater complexity. A paired t-test for difference of means at a significance level of 0.05 yields a difference in mean score is statistically significant ($p<0.001$) between the pre-and post-tests.
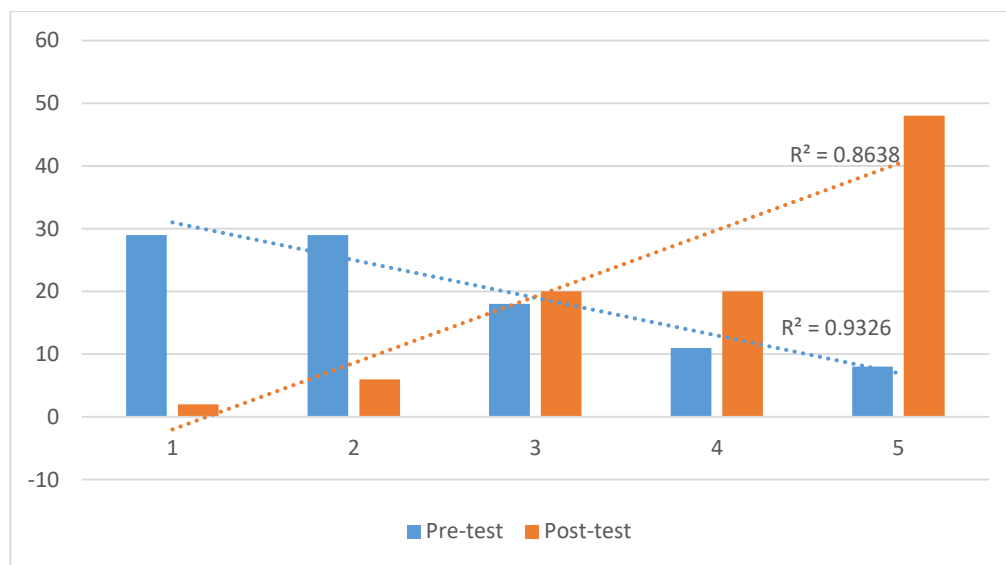

Fig 7. Frequency distribution of structure scores and linear fit for the 2015-16 concept maps utilizing the five scale hierarchy method from STS 4500/4600.

5.0   Discussion

This paper offers evidence about quantity of information generated in the CMs, as well as the shift toward greater complexity in regards to hierarchy and structure in the maps. The findings support a re-thinking of the 'more is better' when evaluating large-scale aggregated datasets of CMs, as students may enter courses with prior assumptions and heuristics that need to be displaced and then replaced with new mental models about the technology-society relationship. There is no way that concept mapping procedures commonly utilized can account for these cognitive and epistemological disruptions. Evaluation of CMs without attentiveness to content present in pre and post- CMs would be misleading for understanding course learning

objectives or ABET criteria (or both). Methodologically, we find a need to further refine how concept map structures are tested for complexity scoring. Presenting two forms of complexity scoring, 4 and 5 scale, indicated the former was not equally suited to handle the complexity in the engineer's concept maps. This points to the need for greater refinement and articulation of map structure as a means to represent the complex and dynamic relationships between technology and society.

The results of the pre-post node comparison between two courses over the past two years shows two distinct outcomes. First, based upon this and previous publications. there is demonstrated knowledge acquisition among the senior students that took the year-long STS 4500/4600 course. Secondarily, there is a high level of variability in the results among the freshman students in the STS 1500 course. This reinforces that the concept maps need to be coded for substance in addition to prescribing to the 'more is better' heuristic in knowledge acquisition. Considering variability differences between freshman and senior students indicates a need to contextualize concept mappers. The challenge becomes what metrics count as more exemplary than others. In a set of courses designed to integrate across disciplinary divides, the richness of concepts and linkages can become overwhelming. This core assessment method in concept mapping continues to offer promising results, but it requires interrogation with larger datasets and more inter-university comparison studies that can include aspects of the student population, institutional environment, and curriculum.

The comparison of the hierarchy and structure coding schema was insightful for a number of reasons. The prior work by Foley et al [7] was with pre-service teachers that have a different skill set. The scoring system in this case offered clear, statistically robust results between the pre-post concept maps. However, the diversity and complexity of the undergraduate engineer's concept maps in the pre-test quickly raised concerns and initiated the creation of an alternative schema that might better account for differences in complexity. The revised, expanded and more nuanced coding structure shows clear changes in the pre-post comparison. However, in broadening the scope of hierarchy interpretability, replicability challenges between coders might arise. This should be tested with inter-rater reliability scoring by additional coders and matched sets of concept maps.

The use of an artifact as prompt appears to be a novel approach in concept mapping. The intent was to initiate mapping from a shared artifact to see how participants may link to social context, including the political and ethical dimensions. While there was no difference between the artifacts in relationship to our analysis of difference, comparative studies between words, phrases, and objects might be an important methodological consideration for future research. The selected artifacts were chosen based upon the ubiquity of the objects and the possibility they encountered these devices in previous courses. However, explorations of alternative objects that are directly connected to technically focused course material (e.g. steel I-beam in a structural engineering course or seatbelt component in a transportation safety course) or indirectly connected (e.g. a baseball in a course about system-level statistics or a hand shovel in a course about subsurface engineering) could offer cognitive prompts that resonate more with engineering students.

Admittedly, there are limitations that we want to acknowledge. First, we do not directly compare words to objects, which is a necessary test to understand how the student-generated concept maps change between words and material objects. While both researchers administered and were present for the duration of the STS 4500/4600 concept mapping activities, they were present for only a few of the CM activities in STS 1500. This brings into question elements of the instruction in mapping and rigorousness of timing the mapping activity despite the use of a script for the course instructors. Finally, it is challenging to document motivation levels for completing the concept mapping activity and the eventual output.

6.0   Conclusion

This paper incrementally advances a few key aspects of concept mapping as a technique to assess knowledge acquisition in engineering education. It probes questions that are fundamental to the design, administration, and analysis of this assessment tool. While limitations are noted and additional research is needed to verify some of the results, we are confident that our approach will support the continued development of this methodological instrument and may yet offer an efficient and effective means to assess learning outcomes in the challenging set of professional skills and systems thinking as we develop a dataset of student knowledge acquisition overtime.

7.0   References

[1]   G. W. Ellis, A. Rudnitsky, and B. Silverstein, "Using concept maps to enhance understanding in engineering education," *International Journal of Engineering Education*, vol. 20, no. 6, pp. 1012–1021, 2004.

[2]   J. C. Nesbit and O. O. Adesope, "Learning with concept and knowledge maps: A meta-analysis," *Review of Educational Research*, vol. 76, no. 3, pp. 413–448, 2006.

[3]   A. J. Cañas, J. D. Novak, and P. Reiska, "How good is my concept map? Am I a good Cmapper?," *Knowledge Management & E-Learning: An International Journal (KM&EL)*, vol. 7, no. 1, pp. 6–19, Mar. 2015.

[4]   M. Besterfield-Sacre, J. Gerchak, M. R. Lyons, L. J. Shuman, and H. Wolfe, "Scoring Concept Maps: An Integrated Rubric for Assessing Engineering Education," *Journal of Engineering Education*, vol. 93, no. 2, pp. 105–115, Apr. 2004.

[5]   S. Ferguson and R. W. Foley, "Assessment of Student Learning Outcomes and ABET Accreditation: A Pilot Study of Fourth-Year Engineering Students using Longitudinal Concept Maps," presented at the 2017 ASEE Annual Conference & Exposition, Columbus, OH, 2019.

[6]   ABET, "Rationale for Revising Criteria 3 and 5," 2016. [Online]. Available: http://www.abet.org/accreditation/accreditation-criteria/accreditation-alerts/rationale-for-revising-criteria-3/. [Accessed: 19-Dec-2016].

[7]   R. W. Foley, L. M. Archambault, A. E. Hale, and H.-K. Dong, "Learning Outcomes in Sustainability Education Among Future Elementary School Teachers," Journal of Education for Sustainable, vol. 11, no. 1, pp. 33–51, Mar. 2017

[8]   I. M. Kinchin, "Concept Mapping as a Learning Tool in Higher Education: A Critical Analysis of Recent Reviews," *The Journal of Continuing Higher Education*, vol. 62, no. 1, pp. 39–49, Jan. 2014.

[9]   N. Derbentseva, F. Safayeni, and A. Cañas J., "Concept maps: Experiments on dynamic thinking," *J. Res. Sci. Teach.*, vol. 44, no. 3, pp. 448–465, Mar. 2007.

[10] V. Gouveia and J. Valadares, "Concept maps and the didactic role of assessment," in *Proc. of the First Int. Conference on Concept Mapping*, Pamplona, Spain, 2004.

[11] R. T. Shankar, D. Mitsova, A. Sapat, and D. J. Terrell, "A Case-Study Approach to Interlink Humanities with Engineering Education," presented at the 2017 ASEE Annual Conference & Exposition, 2019.

[12] B. David and J. Marshall, "Resolving Epistemological Tension in Project-Based Introductory Engineering," presented at the 2017 ASEE Annual Conference & Exposition, 2019.

[13] M. E. Garlock, A. Bhatia, and E. H. Laffey, "Effective Approaches for Teaching STEM-literacy for All Majors: The Example of Resonance," presented at the 2017 ASEE Annual Conference & Exposition, 2019.

[14] J. D. Novak, "Concept mapping: A useful tool for science education," *J. Res. Sci. Teach.*, vol. 27, no. 10, pp. 937–949, Dec. 1990.

[15] J. F. DeFranco, K. Jablokow, M. J. Piovoso, and S. S. Richmond, "Objectively Assessing Concept Maps for Knowledge Integration," *International Journal of Engineering Education*, vol. 31, no. 5, pp. 1288–1298, 2015.

[16] K.-E. Chang, Y.-T. Sung, and S.-F. Chen, "Learning through computer-based concept mapping with scaffolding aid," *Journal of computer assisted learning*, vol. 17, no. 1, pp. 21–33, 2001.

[17] A. G. Briggs, S. K. Morgan, S. K. Sanderson, M. C. Schulting, and L. J. Wieseman, "Tracking the Resolution of Student Misconceptions about the Central Dogma of Molecular Biology," *J Microbiol Biol Educ*, vol. 17, no. 3, pp. 339–350, Dec. 2016.

[18] M. A. Ruiz-Primo, "Examining concept maps as an assessment tool," in *Proceedings of the First International Conference on Concept Mapping*, Pamplona, Spain, 2004, vol. 1, pp. 555–562.

[19] J. Turns, C. J. Atman, and R. Adams, "Concept maps for engineering education: A cognitively motivated tool supporting varied assessment functions," *IEEE Transactions on Education*, vol. 43, no. 2, pp. 164–173, 2000.

[20] J. Segalás, D. Ferrer-Balas, and K. F. Mulder, "Conceptual maps: measuring learning processes of engineering students concerning sustainable development," *European Journal of Engineering Education*, vol. 33, no. 3, pp. 297–306, Jun. 2008.

[21] M. Borrego, C. B. Newswander, L. D. McNair, S. McGinnis, and M. C. Paretti, "Using concept maps to assess interdisciplinary integration of green engineering knowledge," *Advances in Engineering Education*, vol. 1, no. 3, pp. 1–26, 2009.

[22] N. Lourdel, N. Gondran, B. Debray, and C. Brodhag, *Sustainable development cognitive map: A new method of evaluating student understanding*. .

[23] Y. Yin, J. Vanides, M. A. Ruiz-Primo, C. C. Ayala, and R. J. Shavelson, "Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use," *J. Res. Sci. Teach.*, vol. 42, no. 2, pp. 166–184, Feb. 2005.

[24] M. Harsh, M. J. Bernstein, J. Wetmore, S. Cozzens, T. Woodson, and R. Castillo, "Preparing engineers for the challenges of community engagement," *European Journal of Engineering Education*, vol. 42, no. 6, pp. 1154–1173, 2017.

[25] D. C. Shallcross, "Concept Maps for Evaluating Learning of Sustainable Development," *Journal of Education for Sustainable Development*, vol. 10, no. 1, pp. 160–177, Mar. 2016.

[26] R. Pierre-Antoine, S. D. Sheppard, and M. Schar, "Utilizing Concept Maps to Improve Engineering Course Curriculum in Teaching Mechanics," presented at the 2014 ASEE Annual Conference & Exposition, 2014, p. 24.1353.1-24.1353.10.

[27] K. W. Jablokow, J. F. DeFranco, and S. S. Richmond, "A Statistical Study of Concept Mapping Metrics," presented at the 2013 ASEE Annual Conference & Exposition, 2013, p. 23.105.1-23.105.14.

[28] I. M. Kinchin, D. B. Hay, and A. Adams, "How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development," *Educational Research*, vol. 42, no. 1, pp. 43–57, Jan. 2000.

[29] J. R. McClure, B. Sonak, and H. K. Suen, "Concept map assessment of classroom learning: Reliability, validity, and logistical practicality," *J. Res. Sci. Teach.*, vol. 36, no. 4, pp. 475–492, Apr. 1999.

[30] L. T. Schlemer, L. de Greef, and T. S. Harding, "Stories of Change: Faculty in Reflective Dialogues," presented at the 2017 ASEE Annual Conference & Exposition, 2019.

[31] S. Muryanto, "Concept mapping: an interesting and useful learning tool for chemical engineering laboratories," *International Journal of Engineering Education*, vol. 22, no. 5, p. 979, 2006.

[32] R Core Team, *R: A language and environment for statistical computing*. Vienna, Austria, 2018.

[33] R. W. Foley, L. M. Archambault, A. E. Hale, and H.-K. Dong, "Learning Outcomes in Sustainability Education Among Future Elementary School Teachers," *Journal of Education for Sustainable*, vol. 11, no. 1, pp. 33–51, Mar. 2017.