Investigating the Use of Concept Maps and Graph-Based Analysis to Evaluate Learning

Dr. Apurva Patel, University of Texas at Dallas

Apurva Patel is a postdoctoral researcher at the University of Texas at Dallas. He earned his BS in Mechanical Engineering from Clemson University. He also completed his graduate studies at Clemson University, earning his MS in Mechanical Engineering (modeling behavior in function structures) in May 2018, and his PhD in Mechanical Engineering (individual differences in function modeling) in August 2021. Apurva's research interests are human behavior in design, including modeling behavior in function modeling, collaborative work, graph complexity, and prediction using artificial neural networks.

Prof. Joshua D. Summers, University of Texas at Dallas

Joshua D. Summers is Professor of Mechanical Engineering at the University of Texas at Dallas. Dr. Summers earned his Ph.D. from ASU (design automation) and his MS (submarine design) and BS (fluidized bed design) from University of Missouri. He has worked at the Naval Research Laboratory (VR Lab and NCARAI). He was formerly a Professor at Clemson University (2002-2020). Dr. Summers' research has been funded by government, large industry, and small-medium sized enterprises. His areas of interest include collaborative design, knowledge management, and design enabler development with the overall objective of improving design through collaboration and computation.

Mr. Pavan Prasanna Kumar, University of Texas at Dallas

Pavan Kumar is a doctoral candidate at the University of Texas at Dallas focussing on Engineering Identity in undergraduate students. He earned his BE in Mechanical Engineering from VTU, Bangalore, India and MS (Ontology of Engineering Design Activities) in Mechanical Engineering from Clemson University in 2008. Pavan has been a serial entrepreneur and founder of several high-impact organisations for the past decade spearheading the hardware and maker movement in India. His areas of interest are engineering design, prototyping, inclusivity, impact and sustainability.

Shanae Lekeisha Edwards, University of Texas at Dallas

Application of Concept Maps to Measure Effectiveness of Engineering Design Learning Intervention for School Teachers

Author 1, PhD Student
Author 2, PhD Student
Author 3, Postdoctoral Associate
Author 4, Professor

University Affiliation

1 MOTIVATION: MEASURE STUDENT LEARNING

Evaluating changes in learning and understanding is essential to educational research and interventions [1,2]. It serves as a critical parameter to gauge the effectiveness and impact of educational programs [3]. This assessment is not just a valuable tool; it is often the most effective means available to measure and quantify the influence of an intervention. Various researchers across STEM and engineering education have emphasized the significance of qualitatively, quantitatively, and objectively measuring changes in learning outcomes [3–13]. Such measurement not only adds credibility to educational interventions but also enhances their reproducibility and precision, enabling more robust conclusions. To assess the effects of an intervention, it is crucial to start by evaluating participants' prior knowledge [2,9,13–17]. This initial assessment considers various factors influencing individuals' experience and knowledge, encompassing professional expertise [2,18], educational backgrounds [9,19–21], personal factors, and social influences. This holistic evaluation of participants' baseline knowledge provides context necessary to support scaffolded learning.

This study compares pre-intervention and post-intervention knowledge to assess the educational module’s direct impact, excluding prior knowledge and extraneous influences. Such a before-and-after comparison is crucial for determining effectiveness in improving learning outcomes. Additionally, our research includes a retention assessment to evaluate the intervention's long-term effects on knowledge sustainability and its enduring impact on participants. This analysis provides insights into the lasting benefits and applicability of an educational intervention.

One of the tools used in assessing learning and knowledge integration are mental models or mind maps. Mental models are internal representations individuals create to understand and predict the world. These cognitive structures, formed through personal experiences and beliefs, guide behavior, decision-making, and problem-solving [22]. Mental models simplify complex information, aiding adaptation and facilitating team communication [23], especially in design and engineering. Mental models serve as cognitive frameworks enabling individuals to navigate their surroundings, understand elements, and grasp interactions. They are deeply ingrained in our perception of how the world works, especially in physical systems. They represent knowledge about system functioning, components, processes, interconnections, and cause-and-effect relationships [22,23]. Individuals may hold diverse mental models of the same product or artifact, sometimes concurrently.

Mental models are applied at both individual and team levels. Team mental models (TMMs), shared among team members, encompass collective knowledge for task understandings and expectations, resulting in enhanced coordination and adaptability. They are categorized into task, process, and teamwork mental models. Task mental models focus on specific team tasks, such as problem articulation, idea generation, explanations, analyses, and decision-making. Analyzing TMMs reveals insights into managing complex design problems by considering team composition, processes, tasks, context, and competencies. Table 1 provides a list of studies that use mental models in engineering design to measure a parameter of interest in their research.

Table 1: Mental Models in Engineering Design Research

Ref.	Level	Participants	Contribution	Type of Research	Sample Size
[21]	Individuals	Elementary School Students	Conceptual understanding before and after intervention	Mixed-Methods	67
[18]	Individuals	Undergraduate Students	Individual differences while eliciting Mental Models	Qualitative	10
[24]	Individuals	Graduate Students	Effect of functional modeling intervention on Mental Models	Empirical and Quantitative	30
[25]	Team	Undergraduate, Graduate and Professionals	Capturing the content of a team mental model while it is constructed (9 teams)	Mixed-Methods	36
[22]	Team	Professionals	Mental Model sharedness through verbal communication (2 teams)	Mixed-Methods	10
[26]	Both	Undergraduate and Graduate	Impact of team interaction structure on individual and shared mental models (12 teams)	Experimental	36

Measuring learning is essential for evaluating teaching effectiveness and curriculum relevance, allowing educators to identify and address student challenges to meet educational and accreditation standards, particularly in fields with specific outcome requirements like engineering. It aids in guiding student efforts, clarifying misconceptions, and bridging knowledge gaps. Furthermore, it provides employers and stakeholders with critical insights into student competencies, informing hiring and investment decisions. A sample of the methods used to measure understanding and learning by capturing the mental models is presented in Table 2. Along with the methods, the type of participants, the type of research, and a summary of the contributions is also included.

Based on the literature reviewed, concept maps are favored for capturing mental models and assessing learning, due to their effectiveness in analyzing central concepts, relationships, and information flow. Moreover, this approach has been supported both qualitatively and quantitatively in previous studies, focusing on creating mathematically precise models for robust quantitative analysis. Therefore, concept maps are a reasonable choice for assessing student learning or change in knowledge before and after an intervention.

Table 2: Various methods to capture mental models and assess learni

Ref	Method	Participants	Research Type	Contribution
[3]	Revealed Causal Mapping (RCM),	UG	Qualitative	Two methods to accurately capture changes in students' knowledge after intervention
	Structural Assessment (SA)	UG	Quantitative	
[27]	Fuzzy Cognitive Maps (FCM)	General	Mixed Methods	Incorporating stakeholder knowledge into environmental decision-making processes
[25]	Latent Semantic Analysis (LSA)	UG, Grad, Prof	Quantitative	Construction and enactment of team mental models in design teams
	Reflective Practice Analysis (RPA)	UG, Grad, Prof	Qualitative	
[28]	Concept maps (Graph Centrality)	Grad	Mixed Methods	Input of an expert model of text, to output of concept map summarizing the key ideas and relationships in the text

2 CONCEPT MAPS

Concept maps use a graph-based structure to illustrate concepts, attributes, and relationships, serving as a visual framework for understanding individual or team mental models. In concept maps, nodes represent concepts, and edges denote the connections between concepts, facilitating the exploration of inter-concept relations. This method is useful for organizing knowledge in fields like artificial intelligence and engineering design. Furthermore, concept maps are key in developing and enhancing mental models, enriching our comprehension of complex information processing within cognitive structures.

Capturing concept maps at important intervals—before, during, and after an intervention—to form a comprehensive series of bipartite graphs, allows for the creation of a complete knowledge representation of the subject matter. This approach aids in understanding the topic from the participants' perspective, facilitating the adjustment of intervention content to address gaps, highlight critical aspects, and ensure a logical progression of material presented. Such a strategy is crucial in the design of instructional content and curriculum and has an impact on the learning trajectories in constantly evolving fields. Table 3 presents a snapshot of the applications of concept maps in engineering design, along with a summary of the contribution of cited work.

Table 3: Concept graph applications in engineering design

Ref	Type	Application in Engineering Design	Contribution
[29]	Research Article	Proposed scheme for numerical representation of graph structures	Conversion algorithm
[30]	Research Article	Efficient retrieval of assembly models with specific functions	Multi-source semantics information and weighted bipartite graph
[14]	Applied Research with Case Study	Integrating and navigating decision-related knowledge	Representation and navigation of decision-related knowledge
[31]	Quantitative	Evaluating students' knowledge integration	Traditional counting metrics and holistic approach to score and evaluate concept maps
[32]	Applied Research	Graph structures and reasoning mechanisms to the conceptual design process	Specifying functional requirements and the structure of the designed object based on use cases and function graphs.
[33]	Empirical Research Paper	Customer choice preferences in using bipartite network analysis	Customers' consideration and choice behaviors modeling using a two-stage network-based approach

Applications of concept maps can be found in various forms throughout the engineering design process:

- <u>Idea Generation</u>: Engineers use concept maps to brainstorm and organize ideas, fostering creativity and exploration.
- <u>System Understanding</u>: They help engineers grasp complex system interrelationships, identifying areas for improvement and innovation.
- <u>Requirements Analysis</u>: Concept maps organize project requirements, ensuring all aspects are considered.
- <u>Design Optimization</u>: Engineers use them to analyze and optimize designs, considering various parameters, components, and criteria.
- <u>Communication and Collaboration</u>: Concept maps facilitate conveying complex concepts to team members, stakeholders, and clients, aiding collaborative decision-making.
- <u>Knowledge Management</u>: They capture and organize engineering knowledge, preserving best practices and lessons learned for future projects.

2.1 Types of Concept Maps

Analogous to the various uses of concept maps, their representation can also be dependent on the application. Figure 1 and Figure 2 illustrate how the complexity in concept maps can vary, going from a simple representation to a very complex, multi-layered network of information flow. Figure 1 shows two concept maps, a) one for the customer of a swimming pool and b) the other for a lifeguard at that swimming pool. These diagrams are concept maps representing the activities and interactions between users and the designed object, in this case, a swimming pool. [32]

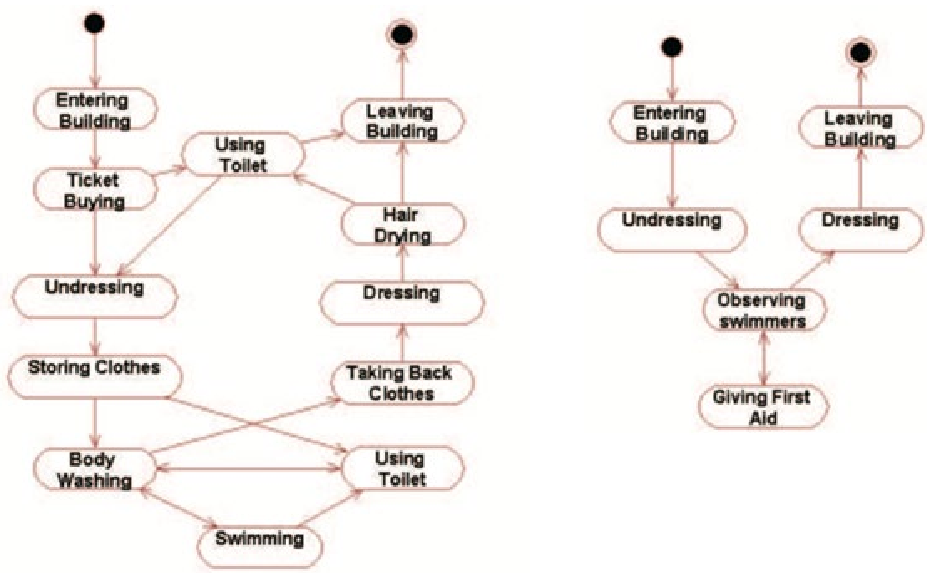


Figure 1: Concept maps for swimming pool customers (left, a) and lifeguards (right, b).

The activity graph for the customer of the swimming pool in Figure 1 (a) shows a sequence of activities that the client is expected to perform, such as entering the building, communicating with staff, buying a ticket, undressing, body washing, and swimming. The graph also includes edges representing the accessibility between different swimming pool areas, such as the entrance area, changing cabins, showers, and the swimming area. A simpler concept graph for the lifeguard is shown in Figure 1 (b), with activities such as observing swimmers, giving first aid, and rescuing swimmers if necessary. These diagrams are concept maps that visually represent the activities and interactions between users and the designed object. They help to identify the functional requirements and constraints of the designed object and can be used to support conceptual design. A more complex concept map is presented in Figure 2 [14], where the yellow ellipses represent the entities in a concept graph, the rectangles indicate the concepts in the concept graph, and the arrows indicate the relationships between these concepts.

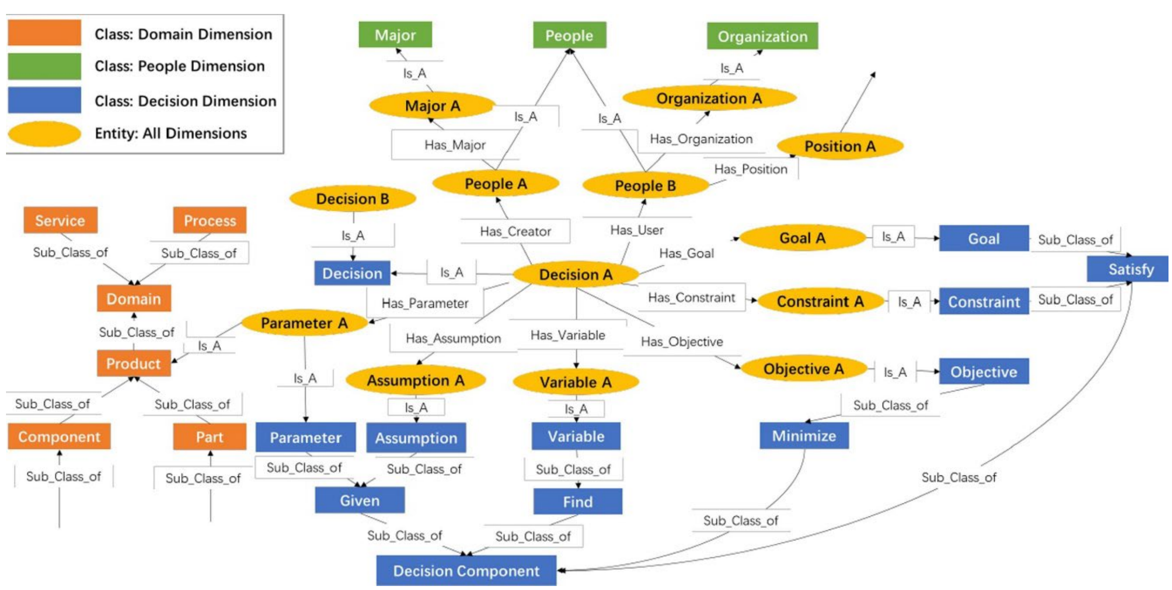


Figure 2: An example of multi-dimensional or multi-plane concept maps.

The concept map in Figure 2 contains three dimensions: decision, domain, and human. These dimensions are interconnected to facilitate the representation and navigation of decision knowledge. The decision dimension contains fundamental concepts representing decision knowledge, including “Given”, “Find”, “Satisfy”, and “Minimize.” These concepts provide perspectives from which decision-makers can seek knowledge. The domain dimension contains concepts specific to a particular domain (e.g. Product, Process, and Service). The human dimension contains concepts that help describe the human aspects and stakeholders involved in decision-making processes. The two different concept maps shown in Figure 1 and Figure 2 represent possibility space for using concept maps to represent understanding of a system.

2.2 Benefits of Concept Maps

Concept maps measure learning by evaluating concept interrelatedness and depth of understanding. When students create concept maps, the structure and connections portrayed offer insights into comprehension and knowledge integration. Maps can be analyzed to assess interrelatedness, knowledge organization, and evolving understanding. Structured rubrics can score concept maps holistically, making them effective for measuring learning. Concept maps serve various purposes in education:

- Knowledge Representation: Maps visually display an individual's knowledge, aiding educators in assessing depth and organization.
- Conceptual Understanding Assessment: They evaluate students' grasp of concepts and connections, gauging accuracy and completeness.
- Misconception Identification: Maps help pinpoint misconceptions or knowledge gaps, guiding targeted instruction.
- Learning Progress Tracking: Maps track learning evolution, comparing maps created at different learning stages.
- Formative Feedback: Educators use maps to provide feedback and identify areas for improvement.
- Metacognitive Reflection: Creating maps prompts metacognition, enhancing self-awareness of learning processes.
- Multidisciplinary Knowledge Integration: Maps assess the ability to integrate knowledge from various disciplines.

Overall, concept maps or concept graphs can be valuable tools for measuring learning by providing a visual representation of students' knowledge, facilitating assessment of conceptual understanding, and supporting the identification of learning progress and areas for improvement. Concept maps have also been used to create journey maps and mental models that capture data at multiple levels, that are detailed/complex and in their dependencies or interconnectedness.

Prior research examines concept maps and the application of graph centrality as a comprehensive metric for evaluating the development of students' mental model structures during the process of writing summaries [28]. The research introduces a model-based methodology for assessing summaries, wherein a concept map is derived from the summary, and various indices are generated from this concept map, spanning multiple dimensions of mental models. Graph centrality is employed as a global metric to capture overarching changes in a student's knowledge structure across these diverse dimensions. The potential of graph centrality is investigated in terms of characterizing the interconnected chain and network structures within a student's concept map.

Concept maps and the utilization of graph centrality are found to offer valuable resources for measuring learning outcomes. They provide a visual representation of students' knowledge, enabling the assessment of conceptual comprehension and facilitating the identification of learning progress and areas needing improvement. Based on the review of concept maps in engineering education, the research objects are discussed next.

3 RESEARCH OBJECTIVES

The objective of this research is to evaluate the use of concept maps as a tool for systematically measuring learning in an engineering design environment. Since concept maps are a largely intuitive tool, students can be asked to create concept maps before and after any lecture or learning activity. This research specifically addresses the *evaluation* aspect of measuring learning through concept maps. It is expected that graph-based complexity metrics and a systematic vocabulary comparison approach can be used to objectively evaluate the change in concept map complexity. More complex concept maps suggest a richer understanding of the phenomena being represented. Moreover, the systematic approach allows for the educators to not only provide a more unbiased assessment, but also lessen the workload of manually evaluating each concept map. Instead, this approach will highlight points of interest and allow for deeper investigation of individual student learning. Assessing changes in learning and understanding is a fundamental aspect of educational research. This assessment not only validates the impact of instructional interventions but also enhances their credibility and precision. By evaluating prior knowledge, measuring post-intervention knowledge, and examining long-term retention, educators gain a comprehensive and insightful understanding of educational programs' effectiveness and enduring value. The pilot study presented in this paper provides a face validity for the approach and identifies areas of improvement.

4 EXPERIMENT: ENGINEERING DESIGN LEARNING INTERVENTION

A controlled user study was designed and conducted to explore whether before-after concept graph generation could be used to capture the changes in mental models, or learning, from an in-class, activity-based intervention. This structure of pre- and post- testing to evaluate the impact of an educational intervention has been used in prior engineering design research to see how a lecture on requirements influences student performance in generating requirements for a design prompt [34]. Further, a third concept graph was collected from participants approximately one month later to capture the retention from the initial intervention. This pilot study is intended to explore the potential for using concept map analysis as a means for assessing learning about engineering design.

4.1 Participants And Experimental Setup

Fifteen participants selected for this study are pre-service teachers participating in a STEM based grant and scholarship program at the University of Texas at Dallas. They participants are between 22 and 30 years old, with fourteen identifying as female and one identifying as male. All are in their final year in the teacher training program. All sessions for this study were held in the same classroom, one with which the participants are familiar, having taken various courses in the room. It is set up with reconfigurable tables that can accommodate six people at each table and a projector/whiteboard at the front of the room. The participants are receiving scholarships as part of the program, but the actual activities are not graded.

4.2 Experimental Context and Procedure

The program includes monthly activities in which the pre-service teachers are introduced to engineering and design. In Spring 2023, a four-month sequence was guided by mechanical engineering faculty in which the pre-service teachers were taught the general engineering design process and several design tools applied to the design and build of small wind tunnels that could be used in their respective classrooms during student teaching. The wind tunnel design and build activity has been used multiple times with elementary students, high school students, and undergraduate engineering students [35,36].

In the Fall 2023 semester, an expanded cohort of participants were introduced to engineering design methods and tools. The participants were taught about design representations in the first session (September). This included an introduction to concept maps, which served to normalize participant understanding of concept maps. One month later, in the next session in October, the first part of the study was conducted. Participants were asked to create a concept graph of engineering design. They were asked to think about how to define engineering design, how to think about engineering design, and what is involved in engineering design. The participants were given five minutes to complete the activity on a blank piece of letter-sized paper. These were collected and the session continued with an introduction of requirements and stakeholders. This served as the intervention for this study which included three parts: a discussion of requirements and analogous terms, a design activity targeting stakeholders and requirements, and a discussion of the relationship between stakeholders and requirements.

The intervention included a high-level discussion on the words typically used in engineering to talk about requirements (desiderata, objectives, goals, wishes, wants, needs, constraints, demands, criteria). A design prompt was provided to the students to address as a team (Figure 3). This prompt was designed to be relevant to the participants as they were all studying to be elementary education teachers and they were being challenged to engage their future students in more STEM related activities.

> Franklin Elementary (3rd grade) teachers want to teach the students about the concept of friction. Specifically, they want to have a “tool” to allow students to test friction with different types of materials. The students should be able to bring their own material and use it on the “tool”. To simplify the concept of the tool, only sliding friction will be tested.
>
> The goal is to design this tool:
>
> - The first step is to identify the stakeholders (post-it notes)
> - The second step is to define the requirements (post-it notes)
> - The third step is to connect the stakeholders and the requirements (post-it notes)

Figure 3: Prompt used to explore stakeholders and requirements.

This activity was followed by discussion about whether all the stakeholders were related to at least one requirement and whether the requirements were connected to at least one stakeholder. The graphs were shared with other teams to see how they were different, capturing and focusing on different aspects of the problem as understood by the pre-service teachers. Next, the teams reconstructed these bi-partite graphs into a relationship table. Differences between the bi-partite graph and the relationship table were discussed from a representation point of view.

Following the intervention, participants were again asked to spend five minutes creating a second concept graph of engineering design using the same instructions. The pre- and post-intervention concept maps were collected for analysis. In the third session held one month later, a third concept graph was collected using the same participants and instructions as the first two graphs. This third graph was collected before any intervention and is used to capture what was retained by the participants.

4.3 Data Collection and Analysis

Data collected in this study was entirely from the participant concept graphs. These were analyzed using a graph complexity approach and based on the vocabulary used to describe the nodes. The concept graphs generated by participants were compared in a within -subject fashion using these two metrics.

4.3.1 Topology-based analysis

The concept maps collected from participants were converted into bipartite graphs to support graph complexity analysis. Each node in the concept map is coded as an entity, and each connection between the nodes is coded as a relation. For this analysis, the directionality of relation was not considered. Similarly, all relations were assumed to have the same weight. An example of a bipartite graph with a corresponding concept map is shown in Figure 4.

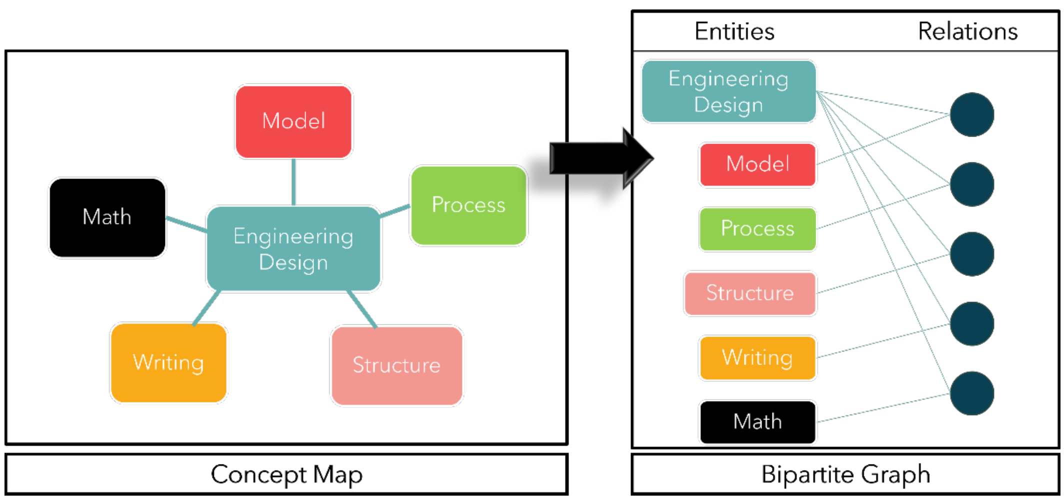


Figure 4: Bipartite graph generated from a concept map

A graph complexity analysis was used to compute twenty-nine complexity metrics for each graph [37,38]. The procedure for calculating the graph complexity metrics is omitted here for brevity; however, the metrics are intended to be a numeric representation of the topological information present in the graph. These metrics are divided into four classes: size, interconnection, centrality, and decomposition. The size metrics relate to the number of elements present in the graph and the number of relations between those elements. The interconnection metrics measure how the different nodes in the graph are connected to each other. The centrality measures provide insight into the clustering and symmetry within the graph. Finally, the decomposition metrics evaluate the solvability and reducibility of the graph. Computational details can be found in prior work focusing on the graph complexity metrics [37,38].

Participant concept graphs are compared topologically using the graph complexity metrics. Three complexity vectors are generated for each participant: $\boldsymbol{C}_b$, $\boldsymbol{C}_a$, $\boldsymbol{C}_r$ corresponding to the complexity metrics for before, after, and retention, respectively. These vectors can be compared in terms of element-wise change in complexity, class-based differences, and Euclidian distance between the vectors. The element wise change in complexity will show how each complexity metric has changed for each individual participant. This will subsequently allow for class-based comparisons where changes in the size, interconnection, centrality, and decomposition of each participant can be evaluated. For positive learning, the size, interconnection, and decomposition metrics are expected to increase. The centrality metrics are not expected to change notably in this study because of the nature of the prompt. The Euclidian distance measures the difference between the complexity vectors as plotted in a 29-dimensional space. This distance comparison will provide insight into how the before, after, and retention concept graphs relate to each other. This within-subject analysis of the data will show whether the effects of the intervention are consistent for all participants. It should be noted that for this pilot study, only the element-wise comparison is presented in this paper.

4.3.2 Vocabulary-based analysis

In addition to the topology, terms used by participants in their concept maps are also analyzed. To support a systematic comparison of the terms, all the terms used by participants in all three iterations of the concept map are collected. The complete list of terms and phrases is processed to remove any duplicate instances and to consolidate similar or highly synonymous terms. This was done primarily using two concepts from natural language processing: stemming and lemmatization.

Stemming reduces a root word to its stem word to normalize sentences for a better understanding [34]. For example, in the dataset, the words ‘testing’ and ‘tests’ would be merged into the common stem ‘test’. Stemming operates on a word without any contextual knowledge and acts independently on each word. As such, a disadvantage of stemming is that it introduces increased polysemy, where a word has multiple meanings [35]. Lemmatization assembles the inflected parts of a word into a single element, or it’s vocabulary form or lemma [34]. It is like stemming but connects multiple words with the same meaning to one word. For example, the words ‘better’ and ‘good’ would be consolidated into ‘good’. Lemmatization is a better approach to generate a primary list from the dataset; however, it is most effective with single words rather than phrases, and more computationally demanding.

Next, the stemmed and lemmatized list of words was reviewed to identify any synonyms. This was done using a similarity DSM approach, where an $n \times n$ matrix is generated with n being list of words. In this matrix, each row or column compares one word in the list to the entire list. As such, the primary diagonal is ignored since these cells refer to comparing a word to itself. Next, the remaining cells are populated by comparing words corresponding to the row and column. Any word pairs that are not similar or the same are populated with zeros. The rest are populated with ones. The matrix is then reviewed to consolidate all similar terms. After processing the terms, a list of 93 unique terms was found. Corresponding to this list, a vocabulary vector is created for each concept graph. Elements of this vector are either “0” or “1” depending on whether the term is present in the concept graph. An example of the vocabulary vector is presented in Table 4.

Table 4: Excerpt from vocabulary coding

Terms	P1	P2	P3
blueprint	0	0	0
brainstorm	0	1	0
clarity	0	0	0
code	0	0	0
collaboration	1	0	1
communication	0	0	1
computer	0	0	0
concept	0	0	1
consult	0	1	0
consumer	0	0	0
cost	0	0	0
create	1	0	0

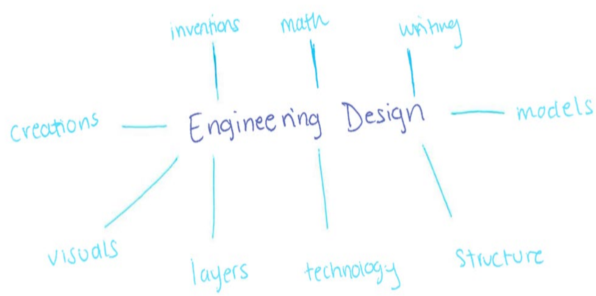
It should be noted that some synonyms were not removed from this list because the context of some words was based on the sub-concept of a node. Other terms were pruned based on the root, for example: the words “creating”, “creation” and “creating new ideas” were all eliminated, and the word “create” was retained. For future studies, a Delphi approach may also be used to further trim the list of terms.

The vocabulary vectors can be compared using various measures such as cosine distance, city-block distance, or hamming distance. Since elements of the vector are binary, Euclidean distance measures should be avoided. In this study, hamming distance is used because it provides a representation of the number of changes necessary to transform one vector into another. Hamming distance is commonly used in error detection where the information is stored in bits (0’s and 1’s). In this case, the hamming distance provides a measure of how many bits needs to be flipped to match the two sets of information. For the vocabulary vectors, this means the number of terms that need to be removed and added. A higher hamming distance will correspond to more changes in the vocabulary, which is expected to correlate with the amount of learning. A more context dependent analysis of the vocabulary vectors could include flagging terms that are highly relevant to the interventions and tracking the change in those terms.

5 RESULTS

This research explored the effectiveness of an educational intervention by measuring immediate learning outcomes and retention over time among participants who were part of a cohort being trained in new concepts for classroom application. Emulating a "train-the-trainer" model, the aim was to enlighten participants about the nature of engineering design, its pedagogical approaches in higher education, and its professional applications. Participants were briefed on the study's scope, making the short-session learning process well-suited for the researchers to observe and quantify the dynamics of learning assessment. This section provides a snapshot of analyzed participant learning presented in a methodical manner which highlights the researchers learning.

While the experiment included fifteen participants, consistent data was collected from only eleven participants. As such, the subsequent analysis and discussion are based on those 11 participants. Among the eleven participants, a variety of concept map complexity was observed. Samples of participant generated concept maps are presented in Figure 5 through Figure 8. Concept maps shown in Figure 5 and Figure 6 are relatively simple, with branching nodes from a single central node and no interconnections between the outer nodes. It is important to highlight that simpler concept maps were not exclusive to the "pre-intervention" phase; they were also observed in post-intervention and retention activities. This indicates that exposure to concept maps during the intervention did not necessarily motivate or equip participants to create more complex concept maps afterwards. Note that Figure 5 and Figure 6 do not show the concept maps from the same participant; they are intended to show the low complexity in both iterations.



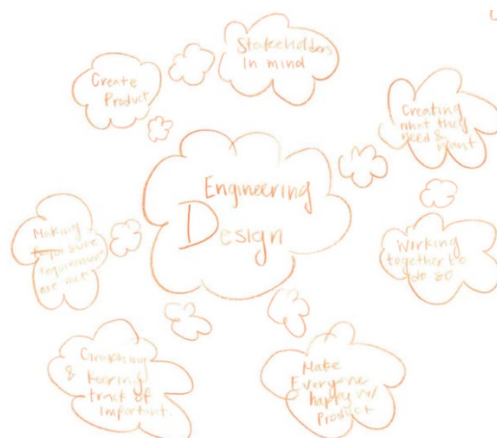
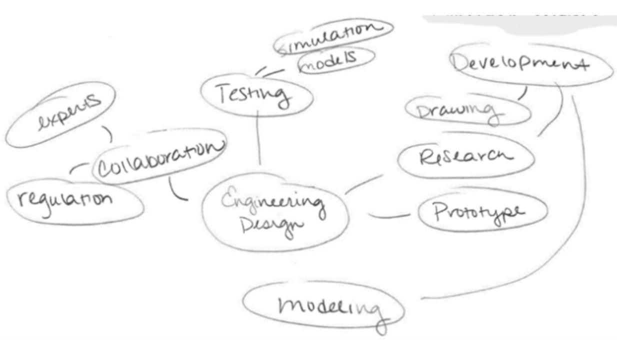


Figure 5:Simple concept map (before) **Figure 6: Simple concept map (after)**

Similarly, more complex concept maps were also observed in all three iterations of concept map generation. These concept maps typically included more than one level of branches emanating out of the central node and/or interconnections between the outer nodes. Figure 7 shows a complex concept map generated by a participant before the intervention where multiple nodes branching out of engineering design produce additional branches themselves. Alternatively, Figure 8 shows a concept map generated by a different participant (after the intervention) where many of the nodes branching out of the central node are connected to each other.



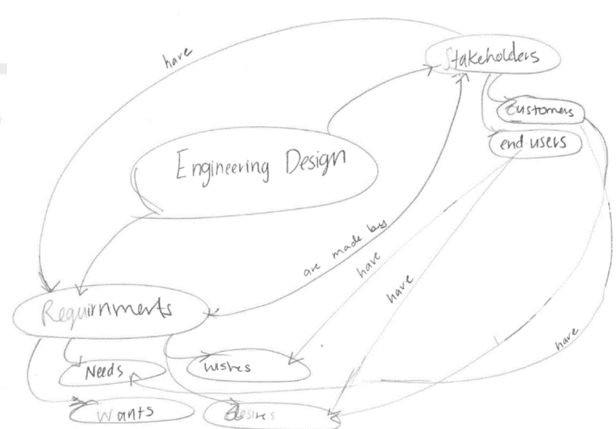


Figure 7:Complex concept map (before) **Figure 8: Complex concept map (after)**

The before, after, and retention sets of concept graphs are topologically analyzed using the complexity metrics. Between the three sessions (before, after, and retention), three pairs of change in complexity metrics are calculated: from before to after, from before to retention, and from after to retention. Since the complexity metrics themselves have varying ranges, a percent change is calculated to support aggregation of complexity metrics. A summary of change in complexity metrics is presented in Table 5, where the "Average" column refers to the average change in complexity metrics for all participants. In this case, it is not only the average of 11 participants, but also of 29 complexity metrics, resulting in an average of 319 values. Similarly, the "Standard Deviation" column shows the average for all participants. This is calculated by first computing the standard deviation of complexity change (n=29) for each participant. Next, an average of these eleven participants is computed to arrive at the number presented in Table 5. The "Minimum" and "Maximum" columns show the smallest and largest change observed for any participant and complexity metric.

Table 5: Statistical summary of change in complexitymetrics

Comparison	Average	Standard Deviation	Minimum	Maximum
Before to After	187%	236%	-74%	5500%
Before to Retention	33%	99%	-100%	1540%
After to Retention	-30%	51%	-100%	252%

Data presented in Table 5 suggests that the largest increase in concept map complexity can be observed when participants create the concept maps after the intervention. The negative change shown for the "After to Retention" comparison suggests that many, if not most, students demonstrated a loss of richness in their concept maps generated a month after the intervention. However, when compared to the concept maps generated before the intervention, those generated in the retention session are still more complex on average. This suggests that some of the knowledge gained during the intervention was retained by the participants. This aligns with the expectations given general understanding of learning and retention. The complexity changes reported in Table 5 provide a general sense of participant learning with respect to the intervention. For a deeper understanding of the phenomena, classes of the complexity metric can be analyzed independently; however, that analysis is not presented here. The purpose of this paper is not to discuss the details of learning induced by the requirement generation intervention, rather to showcase the usability of graph-based analysis to measure learning.

In addition to topological analysis, the vocabulary used by the participants is also compared. In this case, the vocabulary vectors generated for each participant in each session are compared using a hamming distance measure. The hamming distance for each participant in each comparison is presented in Table 6. Unlike the comparison of graph complexity, results from vocabulary comparison are less clear. The vocabulary used in the "after" and "retention" sessions is more similar to each other than those in the "before" session. However, the differences are small in magnitude, suggesting that most of the words used did not change. Further analysis of the vocabulary should include a qualitative comparison of the terms and evaluate their relationship with the intervention. Like the detailed topological analysis, a secondary vocabulary analysis is not presented here. Both remain part of future work, where more data will be collected and the focus of the research is the effects of an intervention, not simply demonstrating the use of this tool.

Table 6: Vocabulary comparison through hamming distance

Participants	Before-After	Before-Retention	After-Retention
P01	0.118	0.118	0.108
P02	0.172	0.172	0.086
P03	0.086	0.097	0.097
P04	0.140	0.108	0.075
P05	0.183	0.108	0.161
P06	0.140	0.151	0.118
P07	0.054	0.108	0.097
P08	0.151	0.118	0.118
P09	0.151	0.161	0.118
P10	0.151	0.161	0.204
P11	0.129	0.151	0.108

6 DISCUSSION AND FUTURE WORK

The pilot study presented in this paper is a precursor to a larger controlled experiment which will include different participants in multiple different activities. As such, no conclusions are provided, instead a discussion of some observations and plans for future work are presented.

Data collected and analyzed in the pilot study suggests that graph-based complexity analysis is a promising approach to measure learning in engineering design. The objective measurements presented in this work show a trend of change in knowledge from before the intervention to after, and ultimately a partial relapse in the retention session. In addition to the systematic evaluations of the concept graphs, the following observations are notable.

- Many of the participants had “requirements” in their concept maps created in the “after” session. Most of them retained that element in the concept maps generated a month later.
- Most of the concept maps generated in the retention session included mode nodes and connection compared to those generated before the intervention.
- There were some terms that were interesting to note: Mistakes, Layers, Stockholders, Blueprint, ‘Efficient and the best’, “Across Fields”, “Pencils and Erasers”, “Design Beauty”, “Crafting”, “Tables”, “Handy”, “Tracker (Airtag)”, and “Purethought”. For an educator, seeing these novel or strange terms can be a starting point for asking questions, and trying to understand how students think these terms fit into the overall understanding of the concept. It provides an opportunity for impromptu formative assessment.
- Some of the written words were difficult to decipher. This may be attributed to a lack of time for drawing and writing, or a desire to finish quickly. This suggests that participants should either be asked to type their responses, or engaged in a post-activity interview where they can not only clarify what they have written, but also explain their reasoning.

The next pilot study will be conducted with more robust data collection. Specifically, the participant pool will be divided into two groups, where one group will be asked to generate concept maps, while the other group will write a short reflection. Moreover, faculty collaborators will be engaged to deploy the learning measurement tool in a classroom environment where students learning can be measured on a weekly basis. This will enable the observation and analysis of knowledge growth over time.

References

[1] Moore, D., and Tananis, C. A., 2009, "Measuring Change in a Short-Term Educational Program Using a Retrospective Pretest Design," *American Journal of Evaluation*, **30**(2), pp. 189–202.

[2] Hill, H. C., Ball, D. L., and Schilling, S. G., 2008, "Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students." *Journal for Research in Mathematics Education*, **30**(4), pp. 372–400

[3] Rutner, P., Armstrong, D., Walker, W., 2004, "The Effectiveness of Two Methods of Capturing Mental Models of Student Learning." *AMCIS 2004 Proceedings*. 397.

[4] Milligan, S. K., and Griffin, P., 2016, "Understanding Learning and Learning Design in MOOCs: A Measurement-Based Interpretation," *Journal of Learning Analytics*, **3**(2), pp. 88–115.

[5] Jonassen, D. H., 1995, "Operationalizing Mental Models: Strategies for Assessing Mental Models to Support Meaningful Learning and Design¬ Supportive Learning Environments." *CSCL '95 Proceedings*. 182-186

[6] Bucciarelli, M., 2007, "How the Construction of Mental Models Improves Learning," *Mind and Society*, pp. 67–89.

[7] Ramalingam, V., Labelle, D., and Wiedenbeck, S., 2004, Self-Efficacy and Mental Models in Learning to Program. *SIGCSE Bull. 36*, **3** (September 2004), 171–175.

[8] Hwang, G. J., Shi, Y. R., and Chu, H. C., 2011, "A Concept Map Approach to Developing Collaborative Mindtools for Context-Aware Ubiquitous Learning," *British Journal of Educational Technology*, **42**(5), pp. 778–789.

[9] Cui, J., and Yu, S., 2019, "Fostering Deeper Learning in a Flipped Classroom: Effects of Knowledge Graphs versus Concept Maps," *British Journal of Educational Technology*, **50**(5), pp. 2308–2328.

[10] Seel, N. M., and Ifenthaler, D., 2012, "Learning-Dependent Progression of Mental Models," *Encyclopedia of the Sciences of Learning*, Springer US, pp. 2032–2036.

[11] Michael, J. A., 2004, "Mental Models and Meaningful Learning," *J Vet Med Educ*, **31**(1), pp. 1–5.

[12] Seel, N. M., 2006, "Mental Models in Learning Situations," *Advances in Psychology*, Elsevier B.V., pp. 85–107.

[13] Isbell, D., and Rogers, J., 2020, "Measuring Implicit and Explicit Learning and Knowledge." *In: The Routledge Handbook of Second Language Acquisition and Language Testing*. Routledge. First Edition.

[14] Hao, J., Zhao, L., Milisavljevic-Syed, J., and Ming, Z., 2021, "Integrating and Navigating Engineering Design Decision-Related Knowledge Using Decision Knowledge Graph," *Advanced Engineering Informatics*, **50**.

[15] Sowa J. F., 1992, “Conceptual graphs as a universal knowledge representation,” *Comput. Math. with Appl.*, **23**(2–5), pp. 75–93.

[16] Mohammed, S., and Dumville, B. C., 2001, “Team Mental Models in a Team Knowledge Framework: Expanding Theory and Measurement across Disciplinary Boundaries,” *J Organ Behav*, **22**(2), pp. 89–106.

[17] Eppler, M. J., 2006, “A Comparison between Concept Maps, Mind Maps, Conceptual Diagrams, and Visual Metaphors as Complementary Tools for Knowledge Construction and Sharing,” *Information Visualization*, **5**(3), pp. 202–210.

[18] Dixon R. A., and Johnson S. D., 2011, “Experts vs. Novices: Differences in How Mental Representations are Used in Engineering Design,” *J. Technol. Educ.*, **23**(1).

[19] Solvie, P. A., and Sungur, E. A., 2006, “Concept Maps/Graphs/Trees/Vines In Education.” *Proceedings of 5th WSEAS International Conference on E-Activities*. Venice, Italy. 220–228

[20] Sands, D., Parker, M., Hedgeland, H., Jordan, S., and Galloway, R., 2018, “Using Concept Inventories to Measure Understanding,” *High Educ Pedagog*, **3**(1), pp. 173–182.

[21] Dankenbring, C., and Capobianco, B. M., 2016, “Examining Elementary School Students’ Mental Models of Sun-Earth Relationships as a Result of Engaging in Engineering Design,” *Int J Sci Math Educ*, **14**(5), pp. 825–845.

[22] Casakin, H., and Badke-Schaub, P., “Measuring Sharedness of Mental Model in Architectural and Engineering Design Teams.” *Proceedings of the 19th International Conference on Engineering Design (ICED13), Design for Harmonies, Vol.7: Human Behaviour in Design*. Seoul, Korea, 169-178

[23] Badke-Schaub, P., Neumann, A., Lauche, K., and Mohammed, S., 2007, “Mental Models in Design Teams: A Valid Approach to Performance in Design Collaboration?” *CoDesign*, **3**(1), pp. 5–20.

[24] Murphy, AR, Banks, HD, Nagel, RL, & Linsey, JS. 2019. "Graduate Students’ Mental Models: An Investigation into the Role of Function in Systems Understanding." *Proceedings of the ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 7: 31st International Conference on Design Theory and Methodology*. Anaheim, California, USA. August 18–21, 2019. V007T06A024. ASME.

[25] Dong, A., Kleinsmann, M. S., and Deken, F., 2013, “Investigating Design Cognition in the Construction and Enactment of Team Mental Models,” *Des Stud*, **34**(1), pp. 1–33.

[26] Wood, M., Chen, P., Fu, K., Cagan, J., and Kotovsky, K., 2014, “The Role of Design Team Interaction Structure on Individual and Shared Mental Models,” *Design Computing and Cognition ’12*, Springer Netherlands, pp. 209–226.

[27] Gray, S. A., Gray, S., Cox, L. J., and Henly-Shepard, S., 2013, “Mental Modeler: A Fuzzy-Logic Cognitive Mapping Modeling Tool for Adaptive Environmental Management,” *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 965–973.

[28] Kim, M. K., and McCarthy, K. S., 2021, “Using Graph Centrality as a Global Index to Assess Students’ Mental Model Structure Development during Summary Writing,” *Educational Technology Research and Development*, **69**(2), pp. 971–1002.

[29] Wyatt, D. F., Wynn, D. C., and John Clarkson, P., 2014, “A Scheme for Numerical Representation of Graph Structures in Engineering Design,” *Journal of Mechanical Design, Transactions of the ASME*, **136**(1).

[30] Han, Z., Mo, R., Yang, H., and Hao, L., 2018, “CAD Assembly Model Retrieval Based on Multi-Source Semantics Information and Weighted Bipartite Graph,” *Comput Ind*, **96**, pp. 54–65.

[31] Besterfield-Sacre, M., Gerchak, J., Lyons, M. R., Shuman, L. J., and Wolfe, H., 2004, “Scoring Concept Maps: An Integrated Rubric for Assessing Engineering Education,” *Journal of Engineering Education*, **93**(2), pp. 105–115.

[32] Szuba, J., Ozimek, A., Schürr, A, 2004. “On Graphs in Conceptual Engineering Design.” *In: Pfaltz, J.L., Nagl, M., Böhlen, B. (eds) Applications of Graph Transformations with Industrial Relevance. AGTIVE 2003. Lecture Notes in Computer Science*, vol 3062. Springer, Berlin, Heidelberg

[33] Fu, JS, Sha, Z, Huang, Y, Wang, M, Fu, Y, & Chen, W. 2017. "Two-Stage Modeling of Customer Choice Preferences in Engineering Design Using Bipartite Network Analysis." *Proceedings of the ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 2A: 43rd Design Automation Conference*. Cleveland, Ohio, USA. August 6–9, 2017. V02AT03A039. ASME

[34] Elena, MV, & Summers, JD. 2019. "Requirement Generation: Lecture Intervention Impact on Variety and Novelty." *Proceedings of the ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 3: 21st International Conference on Advanced Vehicle Technologies; 16th International Conference on Design Education*. Anaheim, California, USA. August 18–21, 2019. V003T04A011. ASME.

[35] Summers, J. D., 2012, “Applied Engineering Service Learning – Design and Build Wind Tunnels for Elementary Classrooms,” *Clemson Collab. Serv. Learn.*, **11–12**, pp. 5–8.

[36] Summers, JD, Joshi, S, & Morkos, B. 2014. "Requirements Evolution: Relating Functional and Non-Functional Requirement Change on Student Project Success." *Proceedings of the ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 3: 16th International Conference on Advanced Vehicle Technologies; 11th International Conference on Design Education; 7th Frontiers in Biomedical Devices*. Buffalo, New York, USA. August 17–20, 2014. V003T04A002. ASME.

[37] Mathieson J. L., Wallace B. A., and Summers J. D., 2013, “Assembly time modelling through connective complexity metrics,” *Int. J. Comput. Integr. Manuf.*, **26**(10), pp. 955–967.

[38] Mathieson J. L., Shanthakumar A., Sen C., Arlitt R., Summers J. D., and Stone R., 2011, “Complexity as a surrogate mapping between function models and market value,” *Proc. ASME Des. Eng. Tech. Conf.*, **9**, pp. 55–64.