Using Concept Maps to Assess Interdisciplinary Integration of Green Engineering Knowledge

MAURA BORREGO
CHAD B. NEWSWANDER
LISA D. McNAIR
SEAN McGINNIS
MARIE C. PARETTI

ABSTRACT

Engineering education, like many fields, has started to explore the benefits of concept maps as an assessment technique for knowledge integration. Because they allow students to graphically link topics and represent complex interconnections among diverse concepts, we argue that concept maps are particularly appropriate for assessing interdisciplinary knowledge integration. The results from a year-long study of a design course in green engineering demonstrate the viability of this approach. However, this research also highlights important issues in faculty scoring of interdisciplinary concept maps that may not be present when maps are used in traditional single-discipline settings. The interdisciplinary setting revealed differences in (1) evaluation criteria, (2) expertise, and (3) investment. We conclude with suggestions for selecting and training scorers to address these issues.

Key Words: assessment, concept maps, green engineering, interdisciplinary

I. INTRODUCTION

Calls for engineers who can operate in interdisciplinary environments have emerged from both government and industry in recent years. In response to the need for skills in interdisciplinary collaboration, multidisciplinary and interdisciplinary design experiences have proliferated in engineering curricula, as even a brief survey of recent scholarship suggests [1, 2]. In engineering education, this is often enacted by bringing students from various disciplines together to work on a group project. As Richter’s content analysis of ASEE conference papers demonstrates, engineering educators are creating innovative experiences at all academic levels and scales that bring students from different disciplines together to address complex technology needs [1]. Smaller efforts involve a limited
number of engineering departments or a specific project. Chen et al., for example, describe a project that brings together students from mechanical, electrical, and industrial engineering to design and build a solar-powered boat [3]. In other cases, universities have well-developed programs supporting a wide range of projects, often with industry sponsors, and typically reaching out beyond engineering [4–7]. Stanford’s Institute of Design, for example, includes business, medicine, the humanities, and education; Howard University brings mechanical and electrical engineering students together with students in marketing and art to collaborate on industry-sponsored projects [8].

What such programs often lack, however, are effective assessment instruments to measure individual student learning gains associated with the interdisciplinary dimension of the projects. Analysis of papers presented at the 2007 ASEE conference reveals that of 86 papers addressing multi- or interdisciplinary design, only two offered methods to assess the multi-/interdisciplinary dimension [1, 2]. To bridge this gap, we focus on one of the hallmarks of interdisciplinary collaboration: knowledge integration. A number of leading researchers have highlighted the ability to synthesize or integrate knowledge and approaches from multiple disciplines as the key factor separating true interdisciplinarity from a multidisciplinary division of labor in which experts work only in their own domains. Interdisciplinary collaborations require individuals from different backgrounds to come together, overcome seemingly incommensurable values and language patterns, integrate knowledge, and reach a high degree of consensus. Good collaborators therefore possess the skills to integrate knowledge and communicate across disciplinary boundaries.

One relatively new assessment technique that can be beneficial in these settings is the use of concept maps. These maps are tools that allow students to articulate knowledge by drawing or outlining core concepts and showing links between ideas (see Figures 1 and 2 as examples). We postulate that concept maps are a particularly appropriate means of evaluating individual interdisciplinary learning and development. In dealing with the complexity of interdisciplinary integration, the multidimensional nature of concept maps can help represent the shifts in understanding that individuals undergo as they exchange knowledge, methods, and values [9]. However, for this tool to be useful, faculty and researchers need an effective approach to scoring concept maps. Besterfield-Sacre and coworkers [10] have developed a reliable holistic scoring method for concept maps used to analyze students’ integration of knowledge within a single discipline. Their study, however, relied on faculty scorers who shared a common discipline, knowledge base, and departmental affiliation. In adapting this approach to an interdisciplinary content area lacking these commonalities, we sought to explore new complexities that might emerge when faculty evaluators from different backgrounds and disciplines come together to apply this holistic scoring technique. Our literature review argues that concept maps are an appropriate means of assessing integration of interdisciplinary content. Consequently, the research question guiding our inquiry is: what factors or concerns arise in faculty scoring of student concept maps when the topic is interdisciplinary?
This study examined concept maps as an assessment tool in an interdisciplinary green engineering design course. The findings both illustrate a successful implementation of this technique and describe patterns associated with the interdisciplinary scoring process. Specifically, differences in evaluation criteria, expertise, and investment all influenced the scoring process. We conclude with a discussion of how to address these factors when selecting and training scorers and implementing the assessment in an interdisciplinary setting.

II. LITERATURE REVIEW

A. Knowledge Integration in Interdisciplinary Settings

The need to train and cultivate individuals who are willing and able to operate in an interdisciplinary environment is a direct result of “the character of problems currently under study, many of which
require the combined efforts of scholars trained in different disciplines" [11]. The complexity and global interconnectedness of many of the engineering challenges of this century require interdisciplinary collaboration to examine the links across fields and to answer crucial questions requiring multiple perspectives [12]. Such interdisciplinary work requires a range of skills addressing both content and process [13, 2]. Most notably, the ability to link and transmit knowledge across unfamiliar contexts is becoming critical. Research suggests that it is essential to develop people “who can serve as both ‘hubs’ and ‘bridges’” between different experts [11]. Thus although many of the descriptions of interdisciplinary design experiences in engineering currently omit measurable, concrete learning outcomes, there is a consensus among scholars that the ability to integrate knowledge, methodologies, and values from multiple disciplines is one of the hallmarks of a good interdisciplinary collaborator [14, 15, 9, 13, 16, 12, 17–19]. It should be noted that although most interdisciplinary collaboration and learning takes place in team settings, we are focusing on an individual’s ability to integrate perspectives and content from different disciplines. Specifically, Boix Mansilla and Duraising identify evidence of this disciplinary integration as a “conceptual framework, graphic representation, model, leading metaphor, complex explanation, or solution to a problem” [19]. This individual ability, as noted earlier, is critical to successful collaboration, and thus provides one indicator of how well individuals will function on interdisciplinary teams. As a graphic representation of the relationship between concepts, concept

Figure 2. Sample Low-Scoring Green Engineering Design Student Concept Map (pre). This was one of several maps tied for lowest score, and faculty commented on its lack of interconnection and heavy reliance on jargon and “buzzwords.”
maps of an interdisciplinary content area would serve as evidence of disciplinary integration (as would design products and projects, which are beyond the scope of this paper.)

Despite this consensus, a number of challenges remain in developing robust techniques to assess this and related outcomes. Accurate assessment of interdisciplinary work is difficult because there is often a lack of “shared strategies and motives” among participants and those who evaluate them [1]. Students, professors, and professionals from different backgrounds have diverse assumptions, values, and epistemic frameworks for how knowledge is accumulated. In addition, participants often bring varying levels of expertise to the collaboration [12]. These differences complicate the assessment process because they result in differences in evaluation criteria, expertise, and professional investment among those responsible for reaching consensus in evaluation. The result is that the scoring process itself becomes an instance of interdisciplinary collaboration.

Current research suggests two key elements that can improve success of interdisciplinary collaboration: increased interaction time and the presence of individuals capable of bridging perspectives. Without a considerable degree of interaction and investment, it is difficult to arrive at shared consensus of what is excellent, average, and poor work (for either collaborators or evaluators). Similarly, the Committee on Facilitating Interdisciplinary Research recommends “extra time for building consensus and for learning of methods, languages, and cultures” in interdisciplinary collaboration [12].

B. Concept Maps in Engineering Education

To address the gap in assessment tools for interdisciplinary learning and collaboration, we argue that concept maps are an appropriate method of assessing the integration of content knowledge from various disciplines into a coherent picture. Concept maps “are graphical tools for organizing and representing knowledge” [20]. They are pictorial essays that allow people to structure knowledge by connecting concepts in a hierarchical and/or linear fashion by identifying concepts and sub-concepts and using cross-links to illustrate the relationships between them. Novak, who has been developing and refining concept maps as a structure-based tool to assess student learning [21, 22, 20], argues that a concept map can be used “not only as a learning tool but also as an evaluation tool, thus encouraging students to use meaningful-mode learning patterns” [20]. Concept maps have been used in programs ranging from middle-school education [23] to statistics [24] to student teacher development [25] to teaching culturally diverse learners [26].

Concept maps have been used to help both engineering and non-engineering students better synthesize knowledge across and between disciplines. Krupczak [27] argues that one way to teach engineering students more effectively is to learn how non-engineering students come to grasp technical concepts. Technological literacy courses have relied on concept maps and a “how things
work” approach to more effectively teach non-engineering students complex, technical ideas. With these methods, students “are receptive and even enthusiastic about learning about technology” [27]. Moreover, the close alignment of interests and expertise levels between first-year engineering students and non-engineers makes this concept map strategy applicable to beginning engineering courses [27]. The approach can help beginning engineering students grasp concepts more easily, introduce them to design issues, “establish appropriate prior knowledge for future courses,” and create a “sense of practical empowerment in novice or even tentative engineers” [27].

Nair has also used concept maps across a broad spectrum of educational levels and disciplines, including secondary and post-secondary students and technical, non-technical, and interdisciplinary topics [28]. For example, as an exercise to understand life cycle analysis and green design, concept maps helped reaffirm the importance of: working in interdisciplinary teams, considering context in problem solving, understanding the design process, making decisions ethically, and solving open-ended problems, and. In addition, Nair notes that a concept map exercise “develops technical context knowledge, interdisciplinary knowledge, decision making skills, and group interaction and communication skills” [28].

In addition to serving as tools to promote student learning, concept maps are also increasingly important in assessing that learning, particularly in engineering education. For example, Walker et al. [29] found that a key difference between students and experts in the engineering design process is the ability to address the broader sociocultural context, especially in terms of ethics and marketing. Students were unable to see beyond the problem statement at hand and address concerns such as motivation for the design, regulatory requirements, or ethics. The researchers used concept maps to evaluate student learning from the beginning to the end of a year-long course based on engineering design. The concept-map approach demonstrated that the student-expert gap closed in two areas: design process and understanding the motivation behind the design.

Similarly, Norman et al. [30] have used concept maps to “directly address how students were able to better synthesize skills developed from multiple courses to new problems” [30]. The course focused on helping students in industrial engineering develop their ability to solve unstructured problems. Concept maps were the primary method used to assess student ability to synthesize knowledge from multiple sources and apply that knowledge to new problems. Moreover, the authors note that the department routinely uses concepts maps for programmatic evaluation to assess knowledge integration in graduating seniors from one cohort to the next.

C. Scoring Concept Maps in Assessment Settings

As Besterfield-Sacre et al. point out, concept maps remain somewhat controversial as assessment tools because of the difficulty of developing a robust approach to map scoring [10]. Approaches
range from counting concepts, links and levels of hierarchy to comparisons of novice and expert maps. In seeking to develop a valid and reliable scoring method appropriate for engineering education, Besterfield-Sacre et al. [10] proposed and validated a holistic assessment method that considers comprehensiveness, organization, and correctness as its evaluation criteria. The approach was tested on a within-discipline analysis of students’ ability to integrate knowledge across a curriculum. Concept maps of industrial engineering were collected from students at multiple points in the curriculum (sophomore, junior, senior); one cohort was followed from sophomore to senior year. Scorers were drawn from departmental faculty, and the results were used to identify and correct programmatic weaknesses. Later, the method was also applied in chemical engineering [31].

Their results demonstrated both the usefulness of concept maps to measure knowledge integration within a discipline and the robustness of the holistic scoring approach. In particular, they found the “holistic rubric is sufficiently sensitive to detect growth in students’ knowledge integration” [10]. Two faculty scorers from the department scored the maps and, where differences arose, discussed the maps to reach consensus. The scoring approach was then translated into a rubric that could be used by other faculty to score the maps. The results demonstrated the validity of the rubric, though they also suggested some unexpected variance across scorers in the organization category that led to an improved final rubric.

Importantly, however, Besterfield-Sacre et al.’s study focused on knowledge integration within a single discipline. The holistic scoring method proved valid where the scorers shared expertise, departmental affiliation, and disciplinary expectations. In contrast, as the literature review demonstrated, in interdisciplinary contexts these shared experiences are notably absent, and in fact pose a significant barrier to the process of reaching consensus.

III. METHODS

To determine the usefulness of concept maps as assessment tools in interdisciplinary settings and to test the holistic scoring process in this context, we applied the method to an interdisciplinary engineering design course in green engineering. Data collection and analysis procedures were designed to directly address the research question: what factors or concerns arise in faculty scoring of student concept maps when the topic is interdisciplinary?

A. Green Engineering as a Site for Interdisciplinary Collaboration and Knowledge Integration

In considering concept maps as a mechanism for assessing knowledge integration, we turned to a green engineering design course as our study site. “Green engineering” and related concepts such
as sustainability are critical nodes of interdisciplinary collaboration because of their prominence in contemporary engineering practice, and the complexity of these fields requires a high degree of knowledge integration.

Green engineering is not a new discipline, but rather a different perspective of traditional engineering practice. It uses the same rigorous application of science and math, but asks disciplines to consider environmental effects in designing products, processes, and systems. The US Environmental Protection Agency (EPA), an early promoter of green engineering, defines it as

the design, commercialization, and use of processes and products, which are feasible and economical while minimizing 1) generation of pollution at the source and 2) risk to human health and the environment. Green engineering embraces the concept that decisions to protect human health and the environment can have the greatest impact and cost effectiveness when applied early to the design and development phase of a process or product [32].

This and similar definitions emphasize applying environmental impacts, including the extraction and processing of raw materials and ultimate disposal of the project, as initial design constraints. Examples of green engineering have been around for decades across disciplines, but the topic has emerged as a distinct approach only in the past ten years. A set of green engineering principles was developed and discussed in detail by Anastas and Zimmerman [33]. A similar list of green engineering principles was developed in May 2003 at the Green Engineering: Defining the Principles Conference held in Sandestin, Florida, through negotiated consensus among stakeholders [34]. Green engineering has also been a topic at the annual American Chemical Society (ACS) Green Chemistry and Engineering Conference for the past twelve years.

European countries are often considered to have successfully implemented more technologies which fit under the principles of green engineering. In the United States, green engineering is being pursued at an increasing rate in industry due to the potential for costs savings, risk reduction, and employee and customer awareness and appeal. The presence of “green” or “sustainable” engineering courses and programs has been growing in recent years. For example, Virginia Tech has offered a concentration in Green Engineering since 2001 [35], and Stevens Institute of Technology [36] and San Jose State University [37] now have minors in the field. Various other schools offer specific courses and programs related to green engineering without a formal degree in this area.

Importantly for this study, not only is green engineering increasingly prominent in engineering curricula, it is also inherently interdisciplinary. The consideration of environmental impacts across
the life cycle of products and processes requires technical consideration of biology, chemistry, and physics in engineering applications. Moreover, there are often non-technical issues such as economics, politics, aesthetics, psychology, and others, that factor in the success of a green engineering approach to a problem. For example, computer scientists, materials engineers, and electrical engineers can collaborate to develop more effective data storage methods that reduce hardware, energy, and resource requirements for computers and data storage facilities. Similarly, computer simulations of experiments or product/process models can reduce or eliminate the need for prototypes in a number of fields.

The ways in which the integration of environmental considerations requires collaboration among diverse disciplines also makes green engineering a popular approach to interdisciplinary design projects in engineering programs in the U.S. and internationally [38–42].

B. Research Setting and Participants

This study involved eleven students enrolled in a two-semester green engineering course; one student dropped the course after the first semester. The course included four students from biological systems engineering, two from materials science and engineering, and one each from management, industrial systems engineering, civil and environmental engineering, and engineering science and mechanics (see Table 1). The course was co-taught by three faculty members whose primary academic backgrounds were chemical engineering/English, English/linguistics, and materials science/chemical engineering.

The course was patterned after a typical capstone design course. To support student enrollment, four departments agreed to allow this course to directly substitute for senior design with no additional requirements. Other departments allowed credit but added limitations including the following:

<table>
<thead>
<tr>
<th>Team #1 (student majors)</th>
<th>Team #2 (student majors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological systems engineering</td>
<td>Biological systems engineering</td>
</tr>
<tr>
<td>Engineering science and mechanics</td>
<td>Biological systems Engineering</td>
</tr>
<tr>
<td>Industrial systems engineering</td>
<td>Biological systems engineering</td>
</tr>
<tr>
<td>Materials science and engineering</td>
<td>Civil engineering</td>
</tr>
<tr>
<td></td>
<td>Marketing</td>
</tr>
<tr>
<td></td>
<td>Materials science and engineering</td>
</tr>
</tbody>
</table>

Table 1. Team composition by student major.
(1) students still needed to go to a department-specific design lecture; and (2) students needed to complete some specific assignments independent of this course. The class met as a group weekly for 75 minutes in a recitation section that covered traditional design topics including the design process, project management, communication, teamwork, and conflict resolution. In addition, the recitation included discussions of green engineering, life cycle analysis, and sustainability. Cradle to Cradle served as the central text for this component [43], and students were expected to both discuss the text in class and apply its principles to their design project.

Two teams were formed based on project interest, as summarized in Table 1. Team #1 (a fifth member dropped after the first semester) had a corporate client interested in reducing the environmental impact of their production waste. Team #2 developed a design for a water filter for the developing world; they had no industry sponsor.

C. Implementation of Concept Maps for Assessment

To evaluate gains in interdisciplinary understanding of green engineering, students were asked to draw two maps, one at the beginning of the course and one at the end of the course. To ensure that the students understood concept maps, an outside expert (a faculty member unaffiliated with the course) introduced the process using the French fry example originally developed by Jennifer Turms and described in Besterfield-Sacre et al. [10]. Students were free to ask questions about how to construct a concept map, but were given full latitude in building their own maps on green engineering design. They were asked to complete a second map at the end of the course, again with the guidance of the outside expert. As reference, students were given a handout during both sessions summarizing the steps to constructing a concept map (Table 2). Consistent with IRB regulations, course instructors were not present for these activities, nor did they find out which students consented to participate in the research until after final grades had been assigned. Eleven students completed the initial exercise independently; one student did not enroll in the second semester and thus did not complete the post-course exercise; that student’s map was dropped from the assessment pool to allow for comparisons of pre and post maps. The maps were stripped of names and information identifying them as pre or post course maps. (Students created the maps using sticky notes on larger pieces of paper; then a graduate assistant recreated the maps in PowerPoint.)

The 20 concept maps were scored by faculty from three departments: materials science and engineering, biological systems engineering, and green engineering. The green engineering scorer was one of the course instructors; just prior to the course, he constructed his own concept map of green engineering. These three faculty were selected as scorers because of several key characteristics. First, all were committed to the place of design in undergraduate education;
one member was the capstone design instructor for his department, another had many years of experience as a capstone project mentor, and the third was a new mentor of capstone projects. All were also committed to providing students with opportunities for interdisciplinary design and had been instrumental in helping the course receive approval as a capstone experience in multiple departments. Finally, all had been consistently supportive of and interested in the concept of green engineering both in their own research and in terms of the institution’s management of its green engineering courses.

The faculty scored the maps in a single two-hour session using the holistic approach developed by Besterfield-Sacre et al. [10] and the rubric shown in Table 3. This rubric required evaluators to score each category on a range of 1 to 3 points; the scores in each category were then summed to determine an overall score for the map. The scoring session involved three components:

1. An engineering education expert “normed” the group by setting expectations for understanding and interpreting concept maps and using examples to bring the group together on scoring values. The expert explained how to use the rubric and answered questions.

<table>
<thead>
<tr>
<th>Steps to Constructing a Concept Map</th>
<th>Adapted from White and Gunstone [44] and <a href="http://www.coff.edu/ete/phl2.html">http://www.coff.edu/ete/phl2.html</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Write down the major terms or concepts you know about a selected topic (e.g., French fries or green engineering design)</td>
<td></td>
</tr>
<tr>
<td>2. Write each concept or term on a separate sticky note.</td>
<td></td>
</tr>
<tr>
<td>3. Sort through the sticky notes, putting terms you do not understand to one side. Also put aside those that are not related to any other term. The sticky notes left over are the ones we will use to construct the concept map.</td>
<td></td>
</tr>
<tr>
<td>4. Arrange the sticky notes so that related terms are close to each other.</td>
<td></td>
</tr>
<tr>
<td>5. Stick the sticky notes to a piece of paper as soon as you are satisfied with the arrangement. Leave a little space for the lines we’ll draw.</td>
<td></td>
</tr>
<tr>
<td>6. Draw lines between the terms you think are related.</td>
<td></td>
</tr>
<tr>
<td>7. Write on each line the nature of the relationship between the terms.</td>
<td></td>
</tr>
<tr>
<td>8. If you put any cards aside in step 3, go back and see if some of them will fit into the concept map you have constructed. If they do, be sure to add the lines and relationships of the new items. Create new sticky notes for new concepts that you think of as you are creating your map.</td>
<td></td>
</tr>
<tr>
<td>9. Your assignment today will be to construct a concept map of “green engineering design.” Write this term at the top or center of your large sheet of paper, then brainstorm related ideas on the sticky notes provided, arrange them, and draw relevant links.</td>
<td></td>
</tr>
</tbody>
</table>

*One Extra Note: Like design, concept mapping is a creative activity with many possible “correct” answers.*

| Table 2. Handout: Steps to Constructing a Concept Map. |
Advances in Engineering Education

Using Concept Maps to Assess Interdisciplinary Integration of Green Engineering Knowledge

2. Each scorer then rated all 20 maps individually, without consulting the other scorers. There was an occasional question to clarify concepts and evaluation practices, but no dialogue about scoring was allowed.

3. The scorers came together to discuss each map and reach a consensus on the scores. The scoring took place in the summer, approximately 2 months after the course ended; none of the faculty scorers were teaching summer school.

D. Data Sources and Methodology

To analyze the scoring process, this study relied on a mixed method approach [45] using a variety of data sources:

1. The concept maps themselves
2. Quantitative individual and consensus scores for each concept map
3. Student information linking pre and post map scores
4. Student majors and faculty scorer disciplinary backgrounds

<table>
<thead>
<tr>
<th>Comprehensiveness—covering completely/broadly</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>The map lacks subject definition; the knowledge is very simple and/or limited. Limited breadth of concepts (i.e., minimal coverage of coursework, little or no mention of employment, and/or lifelong learning). The map barely covers some of the qualities of the subject area</td>
<td>The map has adequate subject definition but knowledge is limited in some areas (i.e., much of the coursework is mentioned but one or two of the main aspects are missing). Map suggests a somewhat narrow understanding of the subject matter</td>
<td>The map completely defines the subject area. The content lacks no more than one extension area (i.e., most of the relevant extension areas including lifelong learning, employment, people, etc. are mentioned)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organization—to arrange by systematic planning and united effort</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>The map is arranged with concepts only linearly connected. There are few (or no) connections within/between the branches. Concepts are not well integrated</td>
<td>The map has adequate organization with some within/between branch connections. Some, but not complete, integration of branches is important. A few feedback loops may exist</td>
<td>The map is well organized with concept integration and the use of feedback loops. Sophisticated branch structure and connectivity</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correctness—conforming to or agree with fact, logic, or known truth</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>The map is naïve and contains misconceptions about the subject area; inappropriate works or terms are used. The map documents are inaccurate understanding of certain subject matter</td>
<td>The map has few subject matter inaccuracies; most links are correct. There may be a few spelling and grammatical errors</td>
<td>The map integrates concepts properly and reflects an accurate understanding of subject matter meaning little or no misconceptions, spelling/grammatical errors</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Concept Map Scoring Rubric, reproduced from [10].
5. Qualitative observation field notes and transcription of the recording from the entire scoring session
This approach applies a *triangulation mixed methods* design. The term “triangulation” in research was first used by Denzin [46] to describe bringing together complementary methods or data sources to offset weaknesses in each (exactly three sources are not required). Applied to mixed methods, data are collected concurrently and combined during the analysis phase, and interpretation involves comparing the results of each to best understand the research question, specifically: what factors or concerns arise in faculty scoring of student concept maps when the topic is interdisciplinary? The quantitative and qualitative components are also assigned equal priority and weighting [47, 45], though the qualitative results require more publication space to describe.

**E. Data Analysis**

The qualitative data was analyzed and integrated with the quantitative findings by iteratively comparing data sets to each other and to the literature. The first three authors listened to the audiotape of the scoring session and developed an initial list of factors associated with scoring interdisciplinary concept maps, supported by the literature on interdisciplinary collaboration. These themes led to hypothesis testing and a final thematic coding scheme for the qualitative data:

1. Evaluation criteria
2. Expertise
3. Investment

Constant comparative method [48] forced us to consider all possible explanations, and to test these against both qualitative data and quantitative relationships. The primary statistical analysis techniques applied to the quantitative data were t-tests comparing pre and post scores of individual students and of groups of students to each other, as well as Cronbach’s alpha for inter-rater reliability among scorers and Pearson correlations to compare individual and consensus scores. After the codes were created and applied to the qualitative data, a second researcher independently checked the codes and applied them to the entire data set. Finally, we shared our findings with the three course instructors and two additional faculty scorers as a final check of our interpretations.

**IV. RESULTS AND DISCUSSION**

Figures 1 through 4 illustrate two pre-course and two post-course maps; Figures 3 and 4 are from the same student. As these figures show, in general the post-course concept maps show an
increase in the number of topics included, the number of appropriate topics included, and the degree of integration among topics. This increased visual complexity suggests an expected increase in knowledge integration as a result of the interdisciplinary experience. The quantitative analysis of the concept map scores demonstrates the student learning gains.

Table 4 presents the pre and post concept map scores for all three individual scorers as well as the consensus scores. It should be noted that for one student (J), the pre score was actually higher than the post test (due to changes in both comprehensiveness and organization subscores). As explained below, the lower post score reflects an inability to adequately address design, while the higher pre-course map score demonstrates more connections and links between concepts. In addition, another student (I) showed no gains between the pre and post-course maps and in fact received the lowest possible score on both maps. Despite these two cases, a paired samples t-test confirms that the student gains were statistically significant (pre: mean = 3.80, S.D. = 1.32, post: mean = 5.10, S.D. = 1.29, t = −2.90, df = 9, p = .018). The distribution of scores is not statistically significantly different from those for junior and senior level students published by Besterfield-Sacre et al.’s [10] Table 7 (t = .470, df = 44, N.S.). The two concept maps for Student A (with the largest gain) are presented in Figures 3 and 4. These overall results provide a foundation for the following three sections, which combine quantitative and qualitative results (where relevant) for each of the three themes.
A. Evaluation Criteria of Faculty from Different Disciplines

Differences in Table 4 seem to reflect a lack of consensus about the structure and boundaries of interdisciplinary content. Similarly, discussion during the scoring session indicated an initial lack of consensus, which required faculty to articulate their interpretation of the criteria. Statistical
At separate times, each instructor observed that they were scoring the maps quite differently from one another. The biological systems engineering scorer posed the question, “Do we agree on anything?” while the green engineering scorer noted, “We don’t have any agreement.” Finally, the materials science and engineering scorer said “We are all over the place.” From the discussion between the three faculty during the consensus scoring, it was clear that they were relying on different, possibly disciplinary, standards to judge the concept maps. There was limited shared expectation of how to apply the criteria to interdisciplinary concept maps, and this was reflected in the faculty scores; Cronbach’s alpha was 0.69. For this measure of inter-rater reliability, 0.70 is considered acceptable, while 0.80 or higher is good.

Analysis of the session transcript suggests that much of this difference may be disciplinary. The biological systems engineering scorer, himself a capstone design instructor in a department that emphasizes the design process, relied heavily on processes feeding back into design to score his concept maps. In one instance, he noted he loved one map because it “feeds back into design” and in general disliked most maps because “everybody has missed taking information over here and having it feed back into the problem statement.” Another criticism he had of the maps was excessive linearity, which he identified as inconsistent with his biology training. In particular, he did not like

<table>
<thead>
<tr>
<th>Student</th>
<th>Pre-course</th>
<th>Post-course</th>
<th>Consensus</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BSE</td>
<td>MSE</td>
<td>GrE</td>
<td>BSE</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>E</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>F</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>H</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>I</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>J</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

**Table 4. Student Concept Map Scores.** Scores from the three criteria (Table 3) are combined. BSE = biological systems engineering faculty scorer, MSE = materials science and engineering faculty scorer, and GrE = green engineering faculty scorer. Students D, F, G and I are biological systems engineering majors, and students B and J are materials science and engineering majors.
when students failed to show the relationship between concepts. He said, “The reason I didn’t like it, maybe because I am in the biological sciences, is that it was all linear.” His standards of scoring were based on the students’ ability to integrate green thinking with design.

While the biological systems engineering scorer relied heavily on integrating process with design, the green engineering scorer noted omissions and inferred what the students knew about green engineering but did not draw. He relied heavily on omissions and gaps in knowledge, particularly when scoring comprehensiveness and correctness. He told the rest of the group to “think of gaps as wrong.” In a similar instance, he said, “There [are] two ways to look at correctness: focusing on the wrong stuff and focusing on how right it is.” To him, for a map to be correct, it must also be comprehensive. Interestingly, however, his scoring was also tempered by his assumption that students possessed core knowledge about green engineering design, but failed to accurately draw that knowledge on a piece of paper. He said, “My assumption was that I think they understood it, but drew it poorly.” On another map he scored, he commented that “they understood it, but drew [it] incorrectly.” This faculty scorer was also one of the course instructors, and may have allowed his familiarity with the class to influence his scoring.

The issue of visualization may explain why the biological systems engineering scorer evaluated the maps differently than the green engineering scorer. Although studies have found that most engineering students at a large public institution prefer a visual learning style [49], the scorers did not agree that these students were able to produce effective visual representations of their knowledge. For example, the faculty member who teaches capstone design in biological systems engineering focused on how the students failed to show the connection between design and process, while the green engineering expert who taught the course assumed that students knew more about green engineering content than they conveyed. In response to a comment from the biological systems engineering scorer about the failure to link process to design, the green engineering instructor noted that “we talked a lot about that [in class].” Also, the actual practice of designing according to green engineering principles was integral to the students’ design projects. Another primary difference between the two raters was that the biological systems engineer focused on organization (connecting concepts), while the green engineering instructor zeroed in on correctness and comprehensiveness. The Cronbach’s alpha for these two faculty scorers was a low 0.56.

The materials science and engineering professor said little about his scoring criteria, but was also concerned about the variance between different students’ abilities to visualize concepts; this factor is discussed further in Section C below. Rather than describing specific scoring criteria, this faculty member expressed a lack of confidence in his green engineering design expertise. Specific comments are included in the following section (B), but it appears that during the consensus phase, this faculty member deferred to the judgment of the two others. In the absence of relevant qualitative
data, we can use the quantitative data to test a hypothesis related to discipline-based evaluation criteria. If this faculty member did not have a clear sense of this interdisciplinary content area, did he perhaps substitute disciplinary criteria? If so, then he might have inadvertently scored students from his own discipline higher than other students. An independent samples t-test reveals that this is indeed the case: the materials science and engineering faculty member scored materials science and engineering students higher than other students ($t = 5.088, df = 18, p = .001$). Neither of the other two faculty scored these students higher than other students (biological systems scorer $t = .916, df = 2.606, N.S.$; green engineering scorer $t = 1.111, df = 18, N.S.$). Table 4 lists all the scores, and readers will note that some of the highest scores were assigned by the materials science faculty member to the materials science students, even though names and majors were stripped from the maps. Ultimately, however, the sample size is too small to draw definitive conclusions. This particular finding suggests interesting directions for future investigations.

B. Expertise in the Interdisciplinary Content Area

Both quantitative and qualitative data demonstrate that the green engineering scorer was viewed as an expert by the other two faculty scorers. The green engineering instructor’s scores were highly correlated to the final consensus scores (Pearson Correlation = .85, $p = .002$). The materials science faculty member’s scores were also statistically significantly—but less strongly—correlated to the consensus scores (Pearson Correlation = .67, $p = .036$). It should be noted that although the green engineering instructor had several years of experience in green engineering, his disciplinary training was also in materials science and engineering. The biological systems faculty member’s concept map scores were not statistically significantly correlated to the consensus scores.

Throughout the consensus process, the faculty scorers deferred heavily to the green engineering scorer both because he was perceived as the content expert and because he was one of the course instructors. They asked the green engineering scorer multiple questions about the structure of the class, how the concept map assignment was given, and technical details about the design and process of green engineering. Green engineering questions ranged from simple ones such as spelling out abbreviations for Green Engineering Principles and Life Cycle Analysis to more complex questions about how a nutrient can be a material.

Furthermore, the green engineering scorer took an active role in trying to shape how the others viewed and scored the concept maps. After the norming session, the green engineering scorer noted that “we will re-norm ourselves when we talk about correctness.” He then proceeded to explain to the other two scorers that assessing correctness can either focus on errors or “how right [the overall map] is.” With this explanation, he referred back to gaps in knowledge as a key scoring criterion for both correctness and comprehensiveness.
While the green engineering scorer was seen by both himself and the other two scorers as the expert, the materials science and engineering scorer reflected doubt about his ability to objectively score the maps. He admitted in response to the green engineer’s criteria for correctness and comprehensiveness that he lacked the knowledge to know whether something was missing or right. He said, “I don’t know if I am familiar enough with the subject area to know if it is comprehensive.” Instead of citing specific criteria about what was comprehensive and correct, he used general criteria in his individual scoring more than references to green engineering expertise. For example, he aligned his judgments with the biological systems faculty member’s insistence on connecting process to design—a more central concept of engineering in general. In fact, the materials science and engineering scorer started to use the biological systems scorer’s logic in making arguments. He noted, “It is hard to follow. It didn’t feed back into design.” Because of his lack of content knowledge about green engineering, the materials science and engineering scorer appeared to rely on the expertise of the two members in the group scoring session and used general principles in his scoring.

C. Investment in Concept Maps and the Interdisciplinary Program

In addition to evaluation criteria and expertise, the theme of investment also emerged from the qualitative data. The commitment, in particular, to the use of concept maps to assess student performance was mixed at the beginning of the discussion. There were statements related to not fully embracing the idea of concept maps as a learning tool or a good evaluation mechanism and references to hurrying the evaluation process because of other commitments. However, at the end of the scoring session there was an increased interest in how concept maps could be used to further student learning in all of the various programs with which these faculty were affiliated. There were even statements related to using concept maps in their own classes. Thus over two hours the scorers evolved from being skeptical about concept maps to wanting to learn more about them and use them as a teaching and assessment tool.

At the start of the discussion, the three faculty members questioned whether concept maps were a viable tool to evaluate student learning. The materials science and engineering scorer was the most skeptical. He argued, “Some people have a hard time reading maps; it doesn’t reflect on their knowledge of geography. They have a hard time looking at two dimensions.” He continued with this argument, noting, “There is just a huge variation in how people do homework. Some people have the ability to organize their thought in two dimensions. That is probably related to their understanding of the subject, but probably not.” The green engineering scorer, in contrast, had mixed feelings about the reliance on concept maps to evaluate student learning. He had his doubts, but also relied on his own map to structure the flow of the class. He said, “I taught the course based on the map I had drawn.” Although he had drawn his own map, he still had doubts of whether students (or even
instructors) could accurately draw the complexity of the design process in green engineering. He pointed out, “It is a very hard concept map to do. It is almost too linked to put on paper well.” In scoring one map, he questioned whether the problem was with the map itself or with the students’ knowledge. He said, “It’s a clear miss. They’re not getting that. So the question is, ‘did they miss it because of the drawing?’ or because “it’s hard to draw[?]”

A second type of investment to consider is investment in the green engineering design course itself. While all three faculty scorers were supportive of the course and asked or answered questions during scoring about how it was run, they did not necessarily translate gaps in student learning to discussion of future improvements for the course. In contrast, Besterfield-Sacre and coworkers [10] reported that departmental faculty scoring their own majors were inspired to restructure curriculum after seeing how industrial engineering majors could not connect statistics to other concepts. In the case of an interdisciplinary program, we would expect faculty to have limited investment due to a high level of department-related disciplinary commitments. The departmental organizational structure of U.S. universities provides a strong disincentive for faculty to work with interdisciplinary programs, including scarce resources and diminished recognition for cross-disciplinary efforts [50]. One mode of discussion that reflected a lack of investment in the project (scoring and/or the design course) was the concern of being late for other commitments, particularly missing a bus to get home for the evening. At one point, one of the scorers had to acquiesce to speed up the consensus process. Noting his dilemma, he said, “I was getting desperate. Just for the record, I really wanted a 3.”

Despite time pressures and initial concerns about the validity of concept maps as an assessment tool, the instructors did show a high degree of interest in using them in their own classes. They repeatedly asked the green engineering instructor about the use of concept maps in class. They peppered him with questions about the instructions, examples he gave in class, how much time the students had, who was present during the activity, who was the outside expert who explained the process to students, and whether he had constructed a map himself. By the end of the discussion, the biological systems engineering scorer said, “I think this would be a nice thing to develop for a teaching aid.” He also asked the green engineering instructor for his personal concept map. With enthusiasm, he said, “I can’t wait to get yours.” At the end, the biological systems and materials science scorers also asked for the instructions and wanted to know if they could keep the printout of the concept maps they scored. Even though there was lack of initial investment in the assessment technique, the scorers spent considerable time and commitment in talking about and examining the maps. Even with time concerns, they rarely accelerated the process of assessing the maps. They attacked each map with vigor and a stringent analysis of what was included and what the students failed to do. As they read them and scored them, there was a general degree of excitement about the possible use of them in the classroom.
Given the centrality of knowledge integration in interdisciplinary environments and the power of concept maps to represent complex knowledge networks, we argue here that concepts maps are a valuable tool for assessing students’ interdisciplinary development. As our literature review demonstrates, prior work on concept maps illustrates their assessment value in settings such as K-12 science learning, teacher education, engineering design, and technological literacy. They have been successfully used to evaluate student development in individual courses as well as across curricula because of their ability to represent not simply content mastery, but connections across content areas. This representational complexity makes them ideal vehicles to evaluate one of the hallmarks of successful interdisciplinary collaboration: knowledge integration.

A comparison of pre and post concept maps of students in a green engineering design course clearly illustrates the ways in which this interdisciplinary design experience increased both the number of topics students were able to associate with green engineering (as measured by comprehensiveness and correctness) and the number of connections they were able to make among those topics (as measured by organization). Importantly, the maps also captured variations in student learning, showing little or no gain for some students, as would be expected in any course.

However, while this study illustrates the use of concept maps as an interdisciplinary assessment tool, it also identifies several core implementation issues that faculty need to address when adopting the tool. Specifically, selecting and training scorers is complicated by three factors:

1. Because any interdisciplinary topic, by its very nature, draws together individuals from different backgrounds, scorers may come with highly divergent evaluation criteria. Such variations can occur even in single discipline assessments, but the qualitative data presented here suggests that interdisciplinarity may exacerbate the issue. Besterfield-Sacre reports that disciplinary faculty in industrial engineering and chemical engineering came to consensus quickly when applying the same rubric to disciplinary student concept maps [51].

2. Differences in background also produce differences in levels of expertise with the interdisciplinary topic, even when scorers share a commitment to or affiliation with that topic. Faculty remain rooted in their own knowledge domains, and may be reluctant to trust their expertise in a more complex interdisciplinary area. This reluctance can lead to a reliance on a single evaluator to arbitrate rather than on a more robust process of negotiation seen in disciplinary settings [52, 51].

3. Scorers in interdisciplinary settings working with a new assessment instrument may have varying levels of investment in both the process and the program. Concept maps, though
gaining in popularity, are still a relatively new assessment tool in engineering education. In addition, interdisciplinary programs often reside outside individual departments and affiliate faculty may lack a sense of ownership or commitment—even when they support the general mission of the program. These two factors can work together to again reduce the robustness of the scoring process; the individual(s) with the most expertise or the most investment can dominate the process and consensus can be reduced to deference to an expert.

Each of these factors can be mitigated when selecting and training scorers through a variety of strategies:

1. While norming is always central to the training of scorers, it is even more crucial in interdisciplinary contexts. As with any interdisciplinary collaboration, upfront discussions of differences can be very productive. Begin scoring sessions by asking the scorers to make their evaluation criteria as explicit as possible, and allow substantial time for faculty to discuss the nature of the criteria used for evaluation (more time than allotted for within-discipline norming sessions). We would suggest, in fact, have scorers produce their own concept maps of the topic as a starting point for the discussion.

2. When selecting scorers, evaluate levels of expertise to insure that the faculty participants are relatively evenly matched in their understanding of the topic. If only one individual is recognized as an expert in the field, he or she may likely dominate the process. Balancing expertise levels insures a more robust consensus process.

3. Build in strategies for insuring that scorers have a stake in the outcomes of the assessment. For example, consider selecting individuals who will be involved in revisions of the course or program, who may teach or advise in the program, or who are responsible for larger assessment efforts that include the interdisciplinary setting. Provide the scorers with feedback on the results of the assessment and the subsequent changes to the course or program that stem from the findings. Invite scorers back in subsequent years to build continuity.

Concept maps, as we have shown here, are robust tools for evaluating knowledge integration in interdisciplinary settings, particularly, as described above, when the process of selecting and training scorers takes disciplinary differences into account.

ACKNOWLEDGEMENTS

The authors wish to thank the National Science Foundation for supporting the green engineering design course and graduate assistant Chad Newswander through grants DUE-0633537 and
EEC-0648439. The views expressed in this paper are those of the authors and do not necessarily represent those of the National Science Foundation. We are grateful to Mary Besterfield-Sacre for helpful discussions about the scoring process, Catherine Amelink and Euan Lindsay for helpful discussions about statistical analysis, Whitney Schaefer for transcription, and the three faculty scorers and ten students who agreed to participate in this study.

REFERENCES


Using Concept Maps to Assess Interdisciplinary Integration of Green Engineering Knowledge

AUTHORS

Maura Borrego is an assistant professor of Engineering Education at Virginia Tech. Dr. Borrego holds an M.S. and Ph.D. in Materials Science and Engineering from Stanford University. Her current research interests center around interdisciplinary collaboration in engineering and engineering education. She was awarded a CAREER grant from NSF to study interdisciplinarity in engineering graduate programs nationwide. Address: McBryde Hall (0218), Blacksburg, Virginia 24061; telephone: (540) 231-9536; email: mborrego@vt.edu

Chad B. Newswander is a Ph.D. Candidate at the Center for Public Administration and Policy at Virginia Tech. His research focuses on organizational theory, governance issues, and public-private partnerships. Address, 104 Draper Road, Blacksburg, VA 24060; email: chadn@vt.edu

Lisa D. McNair is an assistant professor of Engineering Education at Virginia Tech and co-Director of the VT Engineering Communication Center (VTECC). She is co-PI on several NSF-funded projects that explore issues of learning, practicing and teaching interdisciplinarity in university and industry settings. Her primary research interests are communication and collaboration in interdisciplinary and distributed settings, institutional structures that encourage transformational learning, and the possibilities of liberal education in engineering programs. Address: McBryde Hall (0218), Blacksburg, Virginia 24061; telephone: (540) 231-1144; email: lmcnair@vt.edu

Sean McGinnis is the director of the Green Engineering Program at Virginia Tech. He has undergraduate degrees in both Chemical Engineering and Materials Science from the University of Minnesota and a Ph.D. from Stanford University in Materials Science and Engineering. He had 10 years experience in corporate research and development before transitioning his career back to higher education in the area of sustainability. His current educational and research interests focus on the environmental impacts of engineering practice, including life cycle analysis, as well as interdisciplinary education using sustainability as a theme. Address: Holden Hall (0237), Blacksburg, Virginia 24061; telephone: (540)231-1446; email: smcginn@vt.edu

Marie C. Paretti is an assistant professor of Engineering Education at Virginia Tech, where she co-directs the Virginia Tech Engineering Communications Center. Her research focuses on communication in engineering design, interdisciplinary communication and collaboration, and design education. She was awarded a CAREER grant from NSF to study expert teaching practices in capstone design courses nationwide, and is co-PI on several NSF grants to explore interdisciplinary collaboration in engineering design. Address: McBryde Hall (0218), Blacksburg, Virginia 24061; telephone: (540) 231-1812; email: mparetti@vt.edu