International Journal of Engineering Education Vol. 34, No. 2(B), pp. 708–722, 2018
Printed in Great Britain

0949-149X/91 $3.00+0.00

Leveraging Self-Assessment to Encourage Learning Through Reflection on Doing*

JACKSON L. AUTREY, JENNIFER SIEBER, ZAHED SIDDIQUE and FARROKH MISTREE
School of Aerospace and Mechanical Engineering, The University of Oklahoma, Norman Campus, Felgar Hall, Rm. 212; 865 Asp Ave.; Norman, OK 73019, USA. E-mail: farrokh.mistree@ou.edu

We contend that, for the engineer of 2020, the ability to adapt to changing circumstances, technologies, and paradigms will be among the most important competencies to possess. We further contend that students can develop the competency to adapt by engaging in continuous learning through reflection on doing. Additionally, we hypothesize that this critical self-reflection can be implemented in engineering 'Design, Build, and Test' courses. In this paper, we present the 'Learning Statement' (LS) as an instrument for learning through reflection while providing instructors an instrument to assess that learning holistically. We explore a framework for implementation of the LS in a senior-level mechanical engineering design course and a method for evaluating the data collected through its use. Using LS data from 76 students in the Fall 2016 AME4163: Principles of Engineering Design course, we implement a version of the bisecting K-Means algorithm to text mine student LSs for patterns in subject matter frequently written about, changes over the course of a design project, and levels of insight demonstrated. We note that students largely focus in their LSs on keywords linked to principles associated with team formation and management as well as prototype construction and testing. Furthermore, we find that student LSs for assignments strongly correspond to targeted themes. Additionally, we find that LSs assessed as more insightful preferentially focus on areas related to team organization, concept generation, and critically analyzing the design process. We find that text mining analysis of LSs confirm patterns in student learning identified in our earlier work while greatly reducing analysis time. Further, we find that students are challenged by the team-based design process and thereby learn lessons dealing with planning, organizational structure, and delegated responsibility in such structures. Finally, we find that effective self-assessment occurs when students connect their learning to specific future utility.

Keywords: experiential learning; design, build and test; text-mining; learning statements

1. Introduction

We contend that the most important competency for the engineering graduates of 2020 must be the ability to critically reflect on their experiences to produce knowledge continuously, thereby being able to adapt to changing circumstances. The need for this competency is driven by rapid advances in technology, shifting global design paradigms, and an increasing need for engineers able to solve challenges which reach beyond the purely technical. We therefore conclude that this new demand requires new educational strategies and instruments. In this paper, we present an instrument for assessing student learning in engineering design courses which both empowers students to reflect on their experiences and learning while improving the ability of instructors to assess and interpret that learning. We note that foundational to the current paper is a paper published in the 2017 ASEE conference proceedings [1]. In the current paper we focus on the implementation of the instrument and expand on the conclusions drawn in [1].

1.1 Motivation

In an editorial for the *Journal of Mechanical Design*, Farrokh Mistree [2] posits that the primary competency needed for engineering graduates today is the ability to adapt to changing circumstances brought on by new technologies and the challenges of globalization. Further, Mistree highlights how engineering design educators are uniquely positioned to help engineering students to build this competency. What is therefore required is a new emphasis in design education; educators must empower students to learn through reflection on doing. To accomplish this, instructors must augment the goal in design courses from the production of some design artifact to the acquisition of new knowledge and the ability to think critically to generate such knowledge. In our prior work [3], we suggest that many current instructors of engineering design courses (where emphasis is placed on artifacts produced by the design process), using evaluation tools such as report grades and artifact performance, often fail to fully understand (and therefore evaluate) student learning arising from the design experience. We are therefore motivated to *demonstrate how an engineering design, build, and test (DBT) course can be structured around reflection on doing and how novel instruments can both enhance that self-reflection and enable instructors to assess it properly*.

1.2 Pedagogical framework

As our principal motivation is built on the concept

* Accepted 12 January 2018.

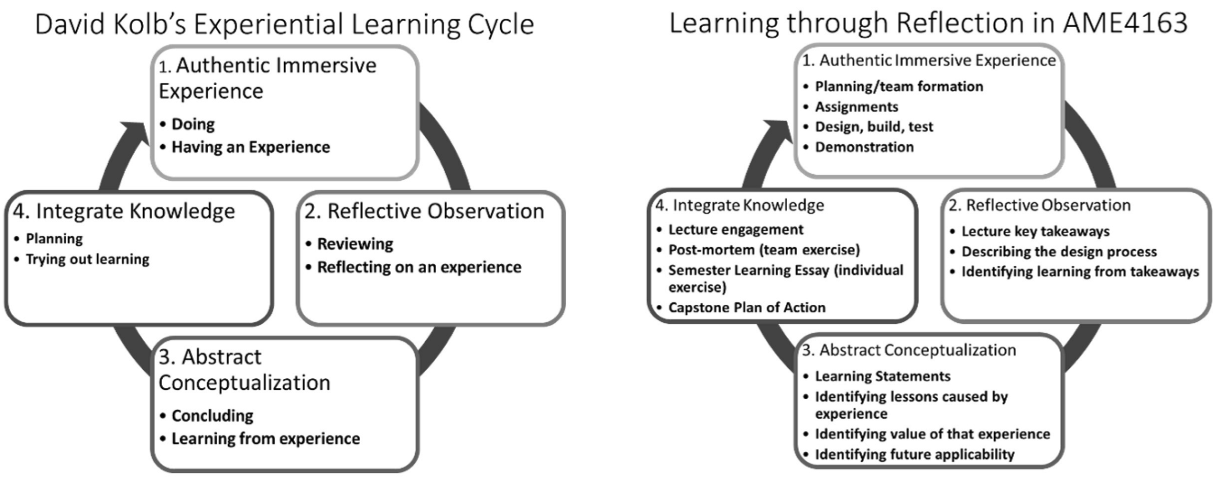


Fig. 1. David Kolb's Experiential Learning and Learning through Reflection in AME4163.

of learning through reflection on doing anchored in authentic, immersive experiences (such as those provided in DBT courses), we largely draw on the work of David Kolb [4], who has identified a mechanism for learning by which a student undergoes some experience, reflects on it, abstracts and then articulates some new knowledge, and then integrates that knowledge into their framework for approaching new problems. The instrument we propose in Section 1.4, the Learning Statement (LS), is anchored in this cycle. Further, as we show in Fig. 1, the structure of our engineering design course, AME4163, is modeled on this cycle.

Of further relevance to our underlying approach is the notion of 'Project-Based Learning (PBL).' According to Dym et al. [5], one of the principal advantages of PBL over traditional educational models is that it facilitates the transfer of old knowledge into new problem contexts, which is one of our central aims for our students. Further, we take a more detailed model for the implementation of this approach in a 'Design, Build, and Test' course from Mistree et al. [6], who have even developed numerous engineering design challenges situated within realistic (though fictional) contexts. We have leveraged one of these problems (and its associated context) in our course.

We suggest that engineering design course instructors who take advantage of the 'Design, Build, and Test' approach and make experiential learning a core part of the course structure will better prepare students to enter industry as junior engineers with the ability to adapt quickly to new circumstances. This assertion has guided our approach to our own course, AME4163: Principles of Engineering Design.

1.3 AME4163 course structure

At the University of Oklahoma, we offer a course in the fall semester preceding the capstone design experience for mechanical engineering seniors titled AME4163: Principles of Engineering Design. Building on the work of Kolb [4] and others, we have made learning through reflection on doing a core component of our course, in which we enable our students to accomplish the following:

1. Internalize the Principles of Engineering Design (POED).
2. Prepare to enter industry as junior engineers following the completion of their undergraduate curriculum.

To that end, we expose students in the course to an authentic, immersive design experience which they must navigate via a structured engineering design process anchored in five Principles of Engineering Design (POED) and their sub-categories, which we outline in Table 1. Over the course of the semester, on teams of four or five, the students develop an understanding of the problem provided, identify customers, classify and prioritize the customers' needs and wants, develop a Requirements List based on those needs and wants, formulate a solution-neutral understanding of their design, ideate concepts to fulfill the identified requirements, compare and contrast the generated concepts using self-established criteria, perform necessary reality checks on the probable modes of concept failure, narrow the concepts to a primary and a backup, realize their primary concept using CAD-modelling software, utilize software to further check the design limitations (component weaknesses, geometric conflicts, etcetera), prototype and test the refined concept design, demonstrate the device as built in a controlled setting in front of their peers, perform a post-demonstration analysis of the design performance, and finally develop a framework for implementing the knowledge obtained from the

Table 1. Principles of Engineering Design Descriptions and Assignment Map

POED	**POED Sub-Category Description**	**Assignment** 1	2	3	4	5
1. Planning a Design Process	1a. Forming a team	✘				
	1b. Accepting and executing a team contract	✘				
	1c. Understanding the problem	✘				
	1d. Proposing a plan of action	✘	✘	✘	✘	
2. Preliminary Design	2a. Ideation: generating concepts		✘			
	2b. Developing concepts (ensure feasibility and realizability)		✘			
	2c. Evaluating concepts; identifying most likely to succeed		✘			
3. Embodiment Design	3a. Refining/modifying most likely to succeed concept			✘	✘	
	3b. Stipulating a Bill of Materials			✘	✘	
	3c. Ensuring functional and technical feasibility, safety, etc.				✘	
4. Prototyping, Testing, and Post-Mortem Analysis	4a. Bill of Materials as built; understand all components					✘
	4b. Ensuring built device meets performance requirements					✘
	4c. Critical analysis of device; causes of success and failure					✘
5. Learning through Doing, Reflecting, and Articulating	5a. Critically evaluating the design, build, and test process					✘
	5b. Articulating internalized POED via learning statements	✘	✘	✘	✘	✘
	5c. Carrying lessons to future: capstone and other ventures					✘

experience into their understanding of design moving forward. These project components are scaffolded into a series of assignments which map to specific, target POED.

We seek to improve the course iteratively, by examining data from the previous year to make adjustments where required. In his thesis, Balmer [7] outlines a model for how these changes are identified and solutions implemented. Further, Balmer describes the target competencies we seek to instill in our students. These competencies are drawn from a variety of sources, including ABET [8], the work of Eggert [9], Lahidji [10], and numerous others [11–14]. In terms of addressing the organizational challenges associated with PBL courses, we draw inspiration for our course structure from the work of Todd [15] and Etlinger [16], who describe the development of two PBL courses designed to instill particular competencies in upper-level engineering design students.

In AME4163, students complete a series of assignments tied to particular POEDs in order to complete their design project. In this paper, we focus only on data collected from Assignments 1–5 and in Table 1 we show how the POED map to each of those assignments. We further scaffold the project through lectures designed to provide student teams with additional context and tools to complete the course and future projects. Although our assignments are scaffolded and require using specific tools to complete them, our lectures allow us to expose students to other useful concepts which they may leverage if they so choose. As our course pedagogy is predicated on emphasizing learning through reflection on doing and deemphasizing the importance of project outcome, we limit the scope of the design project portion of the course to a little over half of the course, with the project demonstration occurring in early November. We set aside the remainder of the course for two major self-reflection exercises, a module on engineering ethics, and an opportunity to plan their approach to their Spring-semester capstone project. We provide the timeline for AME4163 in Fig. 2. The crux of our approach is that, throughout each step of this process, students reflect on doing and articulate their learning in the form of 'learning statements' (LS).

1.4 *The learning statement*

The need for improved assessment instruments for engineering design education has been highlighted by many authors and they cite specifically the importance of self-assessment in such courses. In particular,

Smith et al. [17] have noted improvements in student outcomes when such students engage in critical self-assessment. Others, such as Besterfield-Sacre et al. [18], have expressed a need in engineering design education for new assessment instruments built to capture non-technical skill acquisition. Further, they conclude that self-assessment may be a useful method to accomplish this by accounting for subject student attitudes. Segers and Dochy [19] note that self-assessment may be a useful way to assess student learning in design courses, but with the caveat that students must be trained to critically self-assess. Finally, Olds et al. [20] note that changes in ABET criteria have motivated a need for new forms of assessment in engineering design education, in particular, they call on researchers and educators to collaborate to produce new instruments for such purposes.

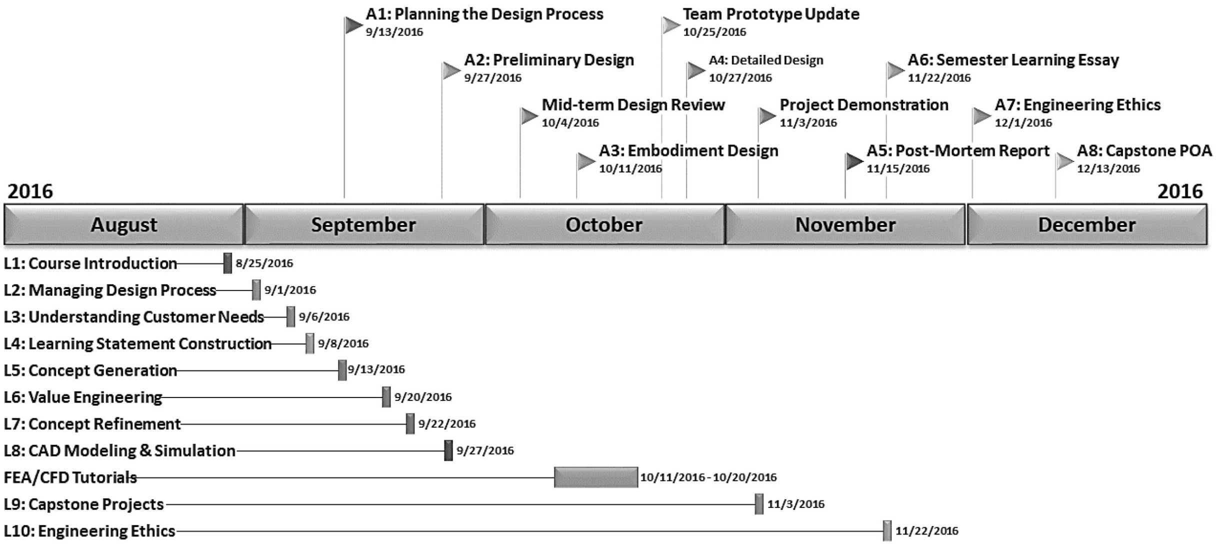


Fig. 2. AME4163 Course Timeline: Lectures, Assignments, and Milestones.

However, while many authors identify the need for improved assessment instruments and have specifically identified self-assessment as a potentially useful avenue for such advances, comparably fewer have investigated self-reflective writing exercises as a possible approach. Among those that have, Turns [21] notes that periodic self-assessment through writing may be useful both to students by providing them a new outlet for instructor feedback and to instructors by providing a resource for improving their courses. Allen et al. [22] find that so-called "learning essays" also improve student reflection on doing while enabling instructors to form an improved picture of learning in their courses.

We therefore propose that engineering design instructors utilize the Learning Statement (LS) to benefit both students, instructors, and educational researchers. Learning statements enable students to freely identify lessons learned as a result of reflecting on authentic, immersive experiences that includes active listening. During AME4163, students write LSs in response to key moments in the course: during lectures, at the end of each assignment, and finally, at the end of the semester in an essay which explores their complete understanding of their learning during the project. We anchor the structure of the LS in David Kolb's experiential learning cycle [4]. As in that construct, students writing LSs reflect on an experience, identify learning from that experience, and then describe the value of integrating that learning in their future endeavors. Although we insist on the structure of the statements, as outlined in Table 2, the students are free to discuss any of their perceived learning as a result of lived experiences and identify the value that learning will provide moving forward. We contend that this exercise is both useful to the student and that the information gleaned from analyzing these statements can be used by instructors to improve their DBT courses. Further, because students submit LSs as individuals and as teams on each assignment, we can produce a picture of the student learning in a team context.

In our previous work [3], we outline a method by which the LSs acquired from students over the course of AME4163 can be analyzed quantitatively. In our prior work, each LS is categorized based on subject matter using the POED sub-categories (see

Table 2. Learning Statement 'Triplet' Structure and Some Suggested Phrasings

Experience *x*	Learning *y*	Value/Utility *z*
Through ***x*** (From ***x***, By doing ***x***, . . .) I did not consider ***x*** initially I thought (expected) ***x*** before/initially	I learned ***y*** I realized ***y*** I found out ***y*** I discovered ***y*** I became conscious of ***y***	Value/Utility ***z*** in future of learning ***y***

Value (Lectures) = Help you transition from a student to a junior engineer and gain insight into how to do the assignments
Value (Assignments) = Principles of Engineering Design

Table 1). For example, a LS exploring the utility of brainstorming in the concept generation phase of a design process would be classified as '2a – Ideation: generating concepts.' Further, we implement a rating scale for quantifying statement 'insight,' which we define as follows:

1. Zero points: Statements earn a rating of zero if the LS is not written to conform to the structure illustrated in Table 2.
   (a) Example: "*Projects tend to be extremely overwhelming when viewed in the holistic sense, but when a plan of attack is proposed that breaks down the project into smaller tasks, the project becomes more conceivable and therefore more manageable.*"—AME4163 student, Fall 2016
   (b) The student fails to put the learning in the context of an experience and therefore is not a LS.
2. One point: Statements receive a rating of one point if the structure is present, but the insight is trivial or obvious.
   (a) Example: "*Through Assignment 1, I have realized that communication protocols are crucial for a team to work together to complete a goal.*"—AME4163 student, Fall 2016
   (b) The student both states something obviously true and neglects to explore any deeper relevance that the learning might have.
3. Two points: Statements receive a rating of two points if in the LS the student demonstrates a connection between their learning and something not explicit to the experience such as a novel circumstance in which the lesson might be applied.
   (a) Example: *"Through considering the customer requirements in greater depth individually, this has taught us more about the entire breadth of the problem and what needs to be taken into account in producing a successful end product, this has a value of allowing us to tailor the device to the end customer more effectively.*"—AME4163 student, Fall 2016.
   (b) The student expresses learning in terms of an experience and then connects that to a future scenario involving a later stage of the design process.
4. Three points: Statements merit a rating of three points if the student exhibits a deeper understanding of the lesson learned and relates its utility to a wider context. Additionally, statements which exhibit clear internalization of any POED merit this rating.
   (a) Example: "*By developing an assembly of our future device, I have learned that preparing a plan of action for assembling it piece by piece in a logical order by stepping through the functions will lead to a better resultant vehicle, more so than just assembling it without regard to the order in which it should be done, which will lead to fewer mistakes in the future when sizing and manufacturing parts and will save our team money and time by eliminating errors and allowing focus to be kept on the completion of the project.*"—AME4163 student, Fall 2016
   (b) The student draws connections while demonstrating a more generalizable lesson learned. The student takes the learning beyond the obvious and directly relates to the fourth POED, which involves manufacturability.

Through the LS instrument, we empower students to simultaneously explore their learning and develop their skills for self-reflection while providing instructors with valuable insight regarding student learning at distinct moments in time. Collecting this data over the course of a design project thereby enables us to form a picture of how student learning is changing in response to new challenges. However, as we have discovered, the two-pronged evaluation method described here is time consuming and subjective. As our course has increased in size, we have been forced to explore new ways to evaluate the LS data. In this paper, we present the case for a text-mining approach to this problem, the method for which we describe in Section 2.2. Given our educational goals for the course, and our contention that the LS may be useful as a research tool for understanding learning in DBT courses, we now seek to outline the specific questions addressed in this paper.

1.5 Research questions and paper outline

In this paper, we observe and reflect on a set of LS data taken from individuals in one section of AME4163 in Fall 2016 in order to identify patterns in the self-reported learning of students. Our sample consists of roughly 1100 LSs collected from 76 students over the five assignments which make up the design project. We seek to understand how the instrument can be used to develop a picture of changes in student learning over the course of a design project. In this paper, we analyze LS data using a bisecting K-means text-mining algorithm to analyze particular subsets of the data for comparison. We also describe the construction of a database to store and facilitate analysis of the more than 10,000 LSs accumulated in Fall 2016 from two sections of AME4163 completing eight assignments and ten lectures.

In this paper, we document how writing learning statements enables students to learn through reflection on doing and how our assessment of these statements can be used to understand student learning in the course better than traditional forms of assessment. Specifically, we address two principal questions in this work:

1. How can the LS be effectively implemented as an instrument for self-assessment in an engineering design, build, and test course? Is text-mining a suitable method for analysis of the data obtained from the use of that instrument?
2. What does text mining analysis of data collected from students in Fall 2016 reveal about how student learning changes in an engineering design, build, and test course over the span of the design project? Does the text mining approach yield new information regarding LSs rated using the scale for 'insight' outlined in Section 1.4?

In Section 1, we have described our motivations for conducting this study and our pedagogical foundations for doing so while also highlighting the need for work involving improved self-assessment instruments and frameworks for their implementation in engineering design courses. Further, we introduce our course, AME4163, a pre-capstone design experience for senior mechanical engineering students, and our chosen self-assessment instrument, the LS, a shortened, discretized version of the 'learning essay' explored by other authors. In Section 2, we explore how we constructed a database to house the LSs collected in Fall 2016 and how this enabled us to both improve our ability to analyze LSs using the two-pronged evaluation method outlined in Section 1.4 and explored in our previous work [3] while also creating an opportunity for the implementation of the text mining approach which forms the basis for this study. In Section 3, we describe the results of our analysis through a histogram and cluster diagram of the most frequently used words in all five assignments, histograms of the most frequently used words for each assignment, and histograms of the most frequently used words in each 'insight' rating category. In Section 4, we discuss the results presented in Section 3 and discuss the implementation of the method in our course, thereby drawing conclusions regarding the potential utility of the approach. In Section 5, we provide additional discussion regarding the applicability of the approach to engineering educators and researchers as well as what work remains to be done.

2. Methods

As discussed in Section 1.4, our prior approach for analysis of LS data is anchored in collecting quantifiable data about each LS and using various forms of standard statistical analysis to identify notable patterns [3]. However, the recent rise of data analytics, text mining in particular, and factors such as the increasing time demand required by our previous approach has led us to pursue alternate means of analyzing the student LSs. We are motivated also by a small but noticeable trend in educational research to utilize text mining to analyze student learning. Researchers such as Frasciello [23] have explored a framework for analyzing engineering student technical writing to identify target writing characteristics. Outside engineering, Wu and Chen [24] highlight the utility of a data mining approach to analyze student performance in online discussions while Kokensparger [25] posits a method for correlating data mined from linguistic features of online writing samples with usage data collected from the same online system. We thus seek to contribute to this research effort by postulating a text mining approach anchored in analysis of the LS self-assessment instrument.

2.1 Database construction

In our prior work with LSs [3], quantitative data were collected by reading each LS submitted by all students and teams, categorizing them based on subject matter and insight, and compiling said categorization for data in relatively simple spreadsheets. This method was useful for the analysis employed in that work, but it was limiting for various reasons. First, the category and insight data were collected in aggregate, which prevented us from looking up ratings for individual students or LSs. Second, obtaining subsets of the data based on particular characteristics (assignment, date, etc.) was painstaking and time-consuming. These and other reasons motivated us to develop a more analysis-friendly approach to data storage and thus the LS database was created.

Data are collected from student assignments and lectures, stored in tab-separated variable (TSV) files, and uploaded to the database. Hosting the database offline and utilizing randomly generated identification numbers allows us to preserve student privacy and avoid potential violations of FERPA. Upon uploading of the TSV files, the database determines whether the LS was written by a team or an individual, and then creates corresponding tables to store the LS based on the attributes associated with each statement. A user interface written in HTML allows us to interact with the database and query particular subsets of the data.

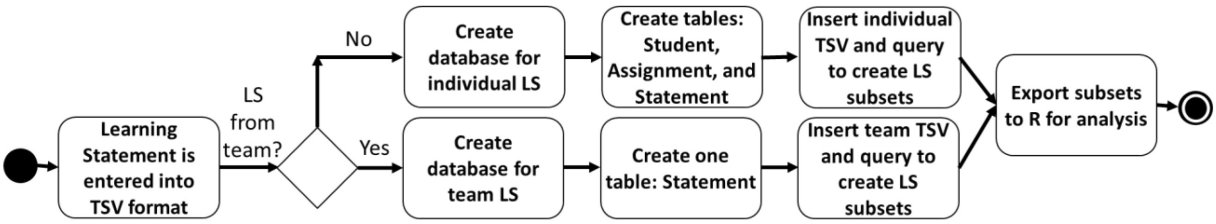


Fig. 3. UML Activity Diagram for the Storage and Querying of LS Subsets.

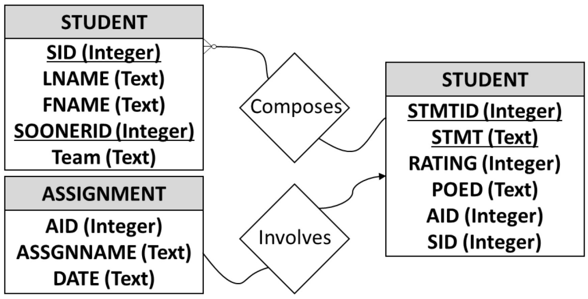


Fig. 4. Entity-Relationship Diagram for Individual LSs Stored in the Database.

We have designed the interface to allow us to obtain subsets of the data as general as all LSs with a particular POED tag or as specific as all LSs from a particular date range from a single student or with a specific rating. In Fig. 3, we diagram this process using a Unified Modeling Language (UML) activity diagram. As we see in the figure, LS text from the queried data subset are output (as a .txt file) for text mining analysis in R. For convenience, the generated data file is stripped of formatting which inhibits the analysis and stored in a pre-specified directory, enabling us to easily locate and access said information to begin the text mining step. Constructed in Python and SQLite, we have designed the database to store LSs in various tables organized around attributes associated with each statement such as a random student identification number, date submitted, assignment or lecture in which the statement was submitted, associated POED sub-category, and insight rating, as we show in entity-relationship diagram in Fig. 4.

2.2 *Text mining*

With the desired subset of queried data stored in our pre-specified directory, we are then able to begin the text mining stage, which takes place entirely within R. In R, the text file is uploaded, and we engage in a series of preprocessing steps. Our first step is to remove all formatting; we remove punctuation, reduce all words to lower case, remove apostrophes, and make other slight changes which might cause the text mining algorithm to stop running correctly within the data set. Next, we remove what are commonly referred to as ‘stop words.’ These are words which we choose to ignore in our analysis because they do not provide useful information about the text. Exactly which words can be removed to improve algorithm performance without sacrificing useful results is a current area of data mining research. Researchers such as Choy [26] have investigated best practices in the context of text mining of Twitter analyses. Accordingly, we use the R package ‘tm,’ which contains the standard English stop words. Additionally, at this stage we are able to specify ‘stop words’ specific to our analysis; words we expect to show up frequently and wish to ignore, such as those we provide students in the LS structure which we outline in Table 2.

Following the pre-processing steps, we are able to set minimum cutoff values for the analyzed words to be considered frequent, so as not to be wholly inundated with data. As our queried subsets may consist of differing numbers of LSs or the frequency of words between different samples may differ sharply, we employ standardized criteria to set the cutoff value. We determine the cutoff value by calculating the mean quantity of words which make up the top one percent of the most frequently used terms in all five assignments. For each assignment (or rating) subset, we then take a number of the most frequent words equal to that mean. Within this set of words for each subset, we set the frequency cutoff equal to the least frequently appearing word, rounded to the nearest factor of five. We are now ready to employ our text mining algorithm.

In this paper, we employ a text mining algorithm called the bisecting K-means clustering technique. As explored by Savaresi and Boley [27], the bisecting

K-means clustering algorithm is an efficient method for hierarchically clustering data based on specified criteria with a guaranteed solution convergence. In particular, Steinbach et al. [28] demonstrate that the approach outperforms other clustering algorithms in text mining applications. In addition, we utilize principal component analysis (PCA) to diagram variance within each data subset. Our chosen PCA method for this study is eigenvalue decomposition, as explored by Jolliffe [29] in his book on PCA methods.

For our analysis, we utilized our text mining algorithm to analyze individual student LSs collected from Assignments 1–5 queried in three primary subsets. In the first, we query all individual LSs from Assignments 1–5 to formulate an overall understanding of the most important student take-aways from the design process. We follow this up with an analysis of individual statements from each of the five assignments queried individually. Finally, we perform our text analysis on three datasets corresponding to Assignment 1–5 LSs queried by 'insight' rating. We explore the results of this analysis in the following section.

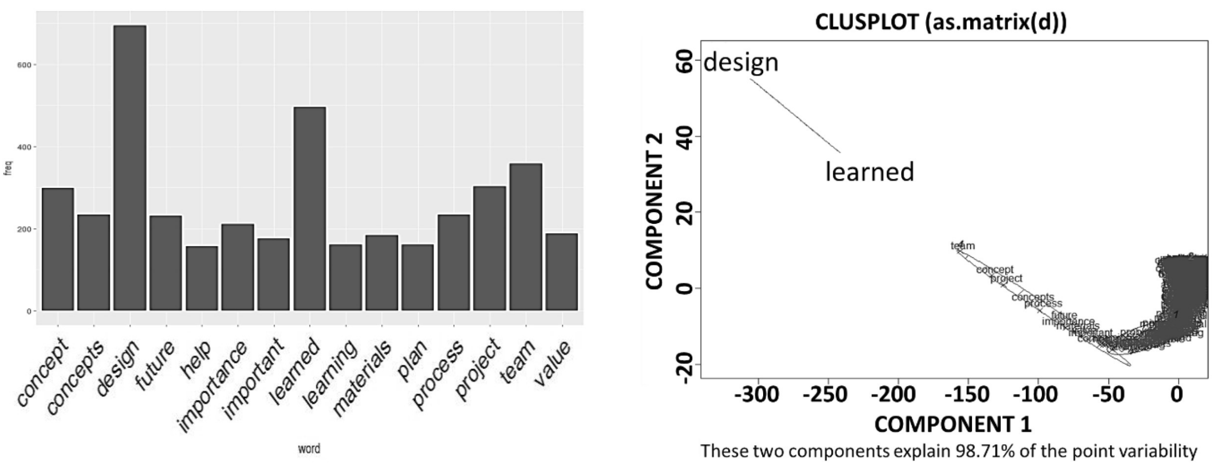
3. Results

In this section, we explore the results from the text mining analysis explained in Section 2.2. Primarily, we focus on word frequency, though we note the utility of exploring the 'proximity' of words frequently used in conjunction with one another. We reserve further analysis along that line for future work.

3.1 *Assignments 1–5 (collectively)*

In the first phase of our analysis, we look at patterns in word frequency generated through text mining for all individual LSs collected from Assignments 1–5. We see in our results, which we show in Fig. 5a and 5b, several notable characteristics about the word choices of students over the course of the design project.

Perhaps unsurprisingly, we see in Fig. 5(a) that the words 'design' and 'learned' are the top two most frequently used words in student LSs over the course of the design project. 'Learned' is an integral part of the LS structure and 'design' is relevant to every POED and appears verbatim in many POED sub-categories (Table 1). More interesting is that the third most frequently used word is 'team' (POED 1a, 1b). We note that many students, in writing their LSs, found at each step of the design process ample reasons to address learning related to team dynamics, organization, and planning, even at late stages of the process. However, if we combine 'concept' and 'concepts' (which may be fair in most instances), then the place of 'team' falls to fourth. Combined, 'concept' and 'concepts' even appear more frequently than 'learned.' Though 'concept/s' might apply to many POED, we find that it is largely associated with POED 2 (generating and evaluating concepts), indicating that this was one of the major areas in which students felt that they learned. Ignoring other words which are arguably relevant to all POED such as 'process,' 'important/importance,' and 'value,' we see in Fig. 5(a) that 'future,' 'materials,' and 'plan' were each well represented. 'Plan' is likely indicative of learning in POED 1, which deals with planning the design process and team formation. 'Materials' most likely refers to POED 3c or 4a, which deal with establishing a planned Bill of Materials and developing a prototype, respectively. We note also in the cluster diagram we present in Fig. 5(b) that 'design' and 'learned' are used so frequently together that



(a) (b)

Fig. 5. (a) Histogram and (b) Cluster Diagram for Aggregated LSs from Assignments 1–5.

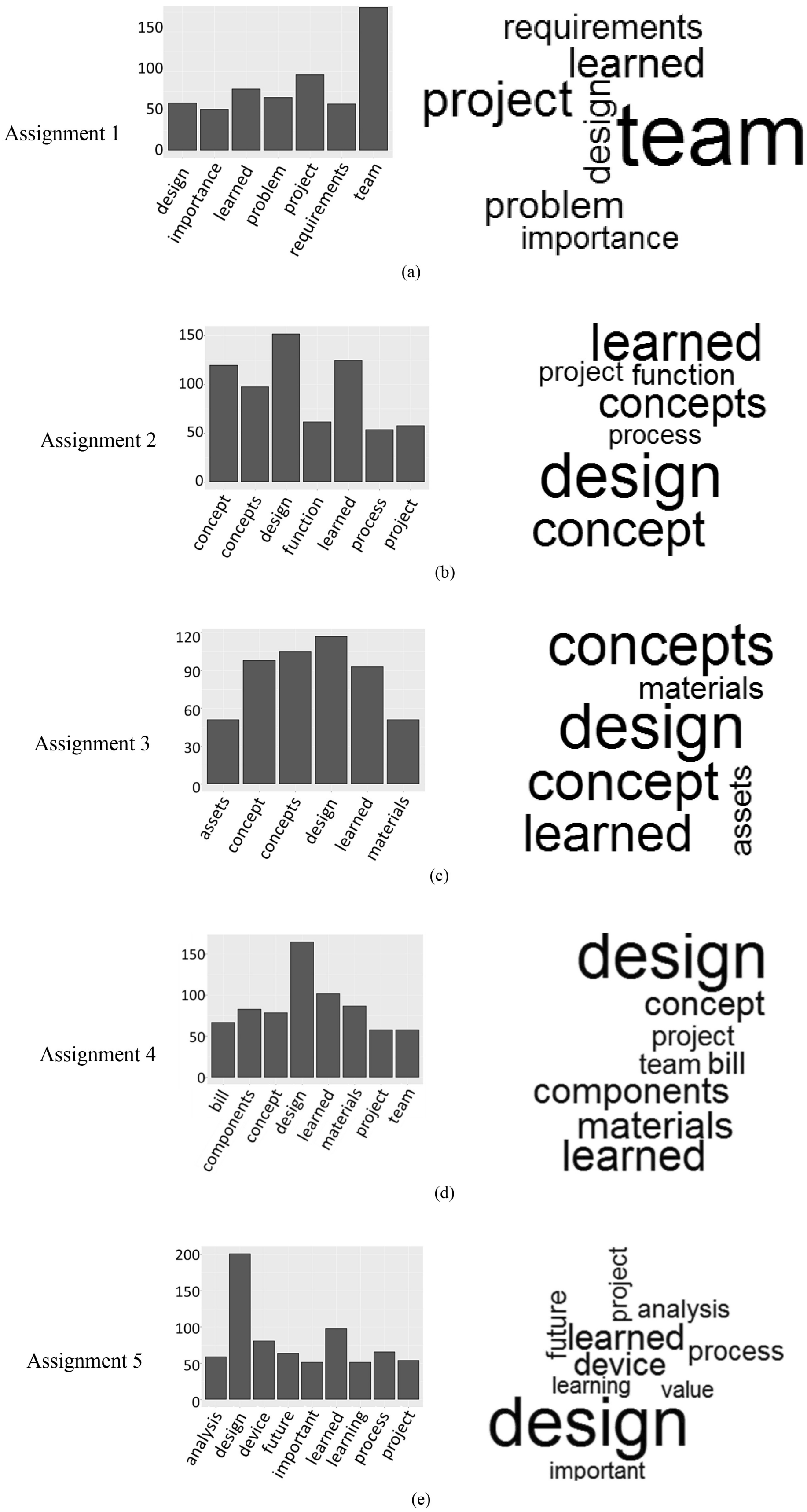


Fig. 6. Histograms and Corresponding Word Clouds of Most Frequently Used Words in LSs for (a) Assignment 1, (b) Assignment 2, (c) Assignment 3, (d) Assignment 4, and (e) Assignment 5

they are in an entirely separate cluster from all other represented words, indicating a lack of notable proximity to any particular words aside from one another. Given that in each assignment, we target specific POED, it is unsurprising that we see that all five POED are generally well represented. To some degree, this implies that students are demonstrating learning in target areas.

3.2 Assignments 1–5 (individually)

In the second phase of our analysis, we examine text mining results for individual LSs queried as subsets broken up by assignment. We see in the results, which we show in Fig. 6 as both histograms of word frequency and word clouds for each assignment, that the most frequently used words in each assignment in many ways map to the target POED established in Table 1.

From Fig. 6(a–e), we see that 'design' and 'learned' are among the most frequently used words in Assignments 1–5. Notably, however, in Assignment 1, which we show in Fig. 6(a), neither are in the top two most used words. Rather, 'team' and, to a lesser extent, 'project' are far more frequently used. As Assignment 1 deals with team formation and planning the design process, we see that notable learning is taking place in these areas. Interestingly, though 'team' is over all five assignments one of the most frequently utilized words, only in Assignment 4, which we show in Fig. 6(d), does it appear in the top terms. In Fig. 6(b) and 6(c) we see that 'concept' and 'concepts' combined are by far the most frequently used words. As 'concept(s)' maps readily to POED 2 and 3, which deal with concept generation and evaluation and concept refinement, respectively, and these two POED are targets in Assignment 2 and 3, we again observe that students are choosing to write about each assignment's target material. In Fig. 6(c) and 6(d), which correspond to Assignment 3 (refining and choosing a primary concept) and Assignment 4 (modifying chosen concept through analysis and preparing for prototyping phase), respectively, we see that students begin frequently using the word 'materials.' Again, this corresponds well to the target assignment POED map we present in Table 1; students begin to choose possible components and 'materials' in Assignment 3 and further refine their selections with greater specificity as a Bill of Materials in Assignment 4. Finally, in Fig. 6(e) we note that words such as 'analysis,' 'future,' and 'process' are among the most frequently used words. Assignment 5 is a post-mortem report, prepared in the weeks following the device demonstration, and in it we encourage the students to critically reflect on both the process they employed and how well they were able to implement it. Though we encourage students to determine the value of newfound learning in their LSs throughout the course, only in Assignment 5 do we see 'future' so well represented in the LSs. Relative to their LSs in other assignments, students completing the post-mortem exercise are thinking more readily about how their knowledge will impact them moving forward..

3.3 Learning between insight rating categories

In the third and final phase of our analysis, we document text mining analysis of LSs from Assignments 1–5 queried by rating using the 'insight' scale we detail in Section 1.4. We present histograms and word clouds of the most frequently appearing words for LS in all three 'insight' categories in Fig. 7.

While in all ratings categories 'design' and 'learned' are among the top three most frequently used words in student LSs, we see in Fig. 7(c) that the frequency of 'learned' compared to other oft-used words is relatively lower for statements rated three. Additionally, comparing LSs by ratings, we also see that variance in the words used shrinks over time. From this we posit that, though all the most frequently used words in LSs rated three, Fig. 7(c), appeared in LSs rated one and two, Fig. 7(a) and (b), only in certain domains were students consistent in writing insightful statements. These areas include LSs tied to POED 1a and 1b ('team'), POED 3a and 4a ('materials'), POED 5a ('process'), and POED 5c ('future'). Subjects conspicuously absent from Fig. 7(c) are those tied to planning, concept generation, and concept evaluation.

4. Discussion

Having presented our data, we now seek to address our research questions, which we posit in Section 1.5.

4.1 Implementation of the LS instrument

In Section 1.1, we discussed our motivation for this work: we have identified that the most important competency that we can be endowing the engineering graduates of 2020 with is the ability to adapt to changing circumstances. We further posit that this is best achieved by equipping students with the ability to critically self-reflect. In Section 1.2, we note that, within engineering design education, there is a need for new assessment instruments which are able to both teach students this skill while enabling educators to more accurately understand student learning in PBL engineering design courses. We thereby posit a course framework to encourage self-reflection in Section 1.3, a self-assessment instrument, the Learning Statement, in Section 1.4, and a data-driven technique which can enable instructors to analyze

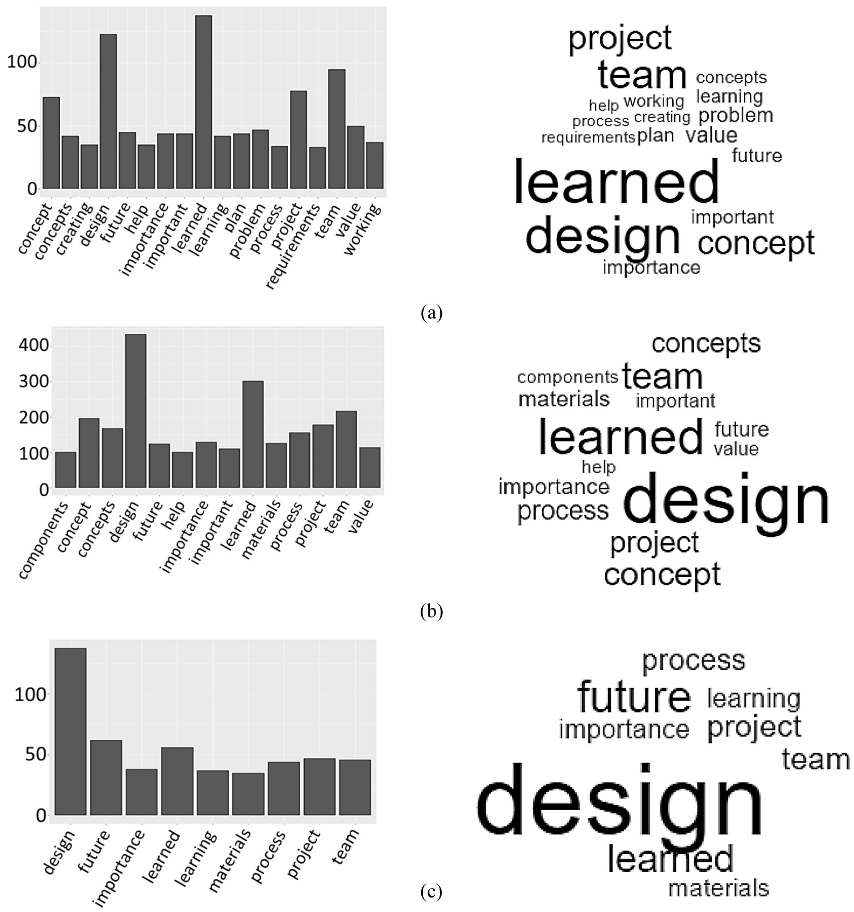


Fig. 7. Histograms and Corresponding Word Clouds of Most Frequently Used Words in LSs (a) Rated One, (b) Rated Two, and (c) Rated Three.

information obtained from that instrument in Section 2. Having thus demonstrated how the text mining analysis of the LS data collected in Fall 2016 from one section of our course, AME4163, can be used to generate a usable picture of how student learning is changing over the course of an authentic, immersive design experience, we feel comfortable concluding that we have satisfactorily answered question one.

From the course framework that encourages constant self-reflection to the instrument used to enable that act and from the construction of the LS database as a research tool which gives us the ability to employ the text mining method, we have demonstrated how information about student learning in our course can be generated in a way which will allow us to make specific course changes, allowing us to make continuous improvements to our course. We can use this information to understand in what areas of our course student learning is strongest and weakest.

4.2 Areas of student learning

By answering research question two, posed in Section 1.5, we hope to understand what we can learn from text mining data about changes in student learning over the course of the design project. We see in Figs. 4 and 5 that there are several notable trends in the areas which students most frequently address in their LSs. From Fig. 5, we find that, overall, students are primarily choosing to write about concept generation/refinement (POED 2, 3a), team formation/organization (POED 1a, 1b), and the design process as a whole (POED 5a). Furthermore, we find that, throughout all assignments, we see in Fig. 5(a) and 6(a–e), 'design' and 'learned' are by far the most frequently used words.

From Fig. 6(a–e), we see in the results evidence that students largely choose to write their LSs about the material targeted in the course assignment, with many of the most frequently used words for each assignment tied to one or more of the target POED for that assignment. Though this may be unsurprising, it confirms that, over the course of a structured design project, students engaging in critical self-reflection throughout the process are finding at least some amount of success in internalizing the target material. This suggests that, even in an open design problem, course targets set by the instructor are impactful on student learning. Of note in particular for this phenomenon might be how readily students appear to connect with the post-mortem exercise (Assignment 5). From Fig. 6(e), we see strong evidence that, looking back on the design process as a whole is prompting students to be more mindful of the impact their newfound learning may have in the future.

Though students appear to largely connect to the target POED over the course of the project, we note some areas in which students do not focus. This has prompted us to make a small change to the assignment scaffold. Starting Fall 2017, we will require that students choose to focus one LS on each target POED for every assignment.

4.3 Degree of student insight

Perhaps the most important takeaway from the results in Fig. 7, in which we show the text mining analysis for LS subsets queried by 'insight' rating, is in what areas student internalization and insight are weakest. That is, while students do identify learning in many areas, for most chosen topics, we suggest that students are not developing the strong connections and insight we want them to demonstrate. In fact, for LSs rated three, we see in Fig. 7(c) that relatively few topics are represented. Among the topics explored in highly rated LSs (rating of three), students are consistently demonstrating strong insight regarding team formation and organization, identifying potential and final materials to realize their designs, and critically evaluating the design process as a whole. In contrast, though students write quite frequently about concept generation and refinement, developing a plan of action, and prototyping, the LSs students are writing dealing with these subjects are relatively weak. This information allows us to answer the second part of research question two, posed in Section 1.5, which deals with whether text mining provides useful information about LSs rated using our 'insight' scale. Given that this approach has allowed us to identify areas in which students are not developing the degree of insight we hope to see and that, consequently, we can make targeted changes to the way the course is taught, this method of analysis is validated.

5. Relevance and future work

Having thus far demonstrated a self-assessment instrument designed to encourage student reflection in engineering design, build, and test courses as well as a useful framework for implementing said instrument, collecting data from it, and analyzing that data to produce new knowledge about our course, we now seek to address the potential utility of this work to engineering design educators and researchers as well as identify further areas of inquiry.

5.1 Relevance to educators

As engineering educators, our goal must be to strive to best prepare students for the conditions we anticipate that they will face when they enter industry as junior engineers. For the better part of two decades, educators have increasingly sought to accomplish this goal via project-based learning. We suggest that by exposing engineering students to an authentic, immersive design experience with realistic constraints and objectives, students are empowered to develop competencies beyond the technical skills traditionally associated with engineering practice. In our work, we have sought to improve upon this approach via an improved framework and an assessment instrument which, when combined, can help improve student ability to critically self-reflect, a necessary step in becoming a continuous, life-long learner. Furthermore, we have demonstrated how deploying such an instrument and analyzing the data collected from its use can be used by educators to make improvements to their DBT courses. We invite other educators to experiment further with the strategies and instrument we have developed. To that end, all material explored in this paper is available upon request, including the open-source software tools we have developed to accomplish our research aims.

5.2 Relevance to researchers

As is true for educators, engineering education researchers are in an excellent position to positively impact the junior engineer of 2020. We note in Section 1.2 that others have identified gaps in pedagogical knowledge pertaining to assessment of student learning in engineering design. In this paper, we address some of the gaps, particularly those dealing with self-assessment and the increasing urgency to shift focus in design courses from a traditional framework in which the importance of student design outputs such as the design artifact is emphasized to a framework in which learning through reflection on doing is encouraged. As noted by Olds et al. [20], it has never been more imperative that researchers and educators collaborate to produce new knowledge for the improvement of engineering design education and it is in that spirit that we make the same offer to engineering researchers as we made to educators: we encourage all to leverage the course framework, assessment instrument, and software tools we present in this paper for continued scholarship in this arena. We hope to see others picking up where we have left off in the near future.

5.3 Way forward

The analysis discussed in this paper covers only the assignments which encompass the span of the design project, concluding with the LSs obtained from the students' post-mortem reports. An additional assignment, the Semester Learning Essay, is completed in the month following the conclusion of the design project. In it, we task students with identifying their most impactful or important lessons learned over the course of the semester. Though similar in principle to the post-mortem report, this exercise is designed to address learning only. One question we will be addressing in the near future is whether the learning expressed in the Semester Learning Essay confirms the patterns we note taking place during the design project or whether the 'distance' between said assignment and the conclusion of the project has provided students with revised perspective.

In addition, we seek to address questions of repeatability in the analysis presented here. Data presented in this work comprise only half of the total student data collected from Fall 2016 (one of two sections). Given that the course was co-taught to both sections by the same two instructors and teaching assistant with no significant changes to the course structure or evaluation methods employed, we hope to address whether the patterns we observe are consistent among two different sampled groups. Collecting data from following years will also provide us the opportunity to address questions of efficacy related to changes made to the course. We hope to address, for example, whether the changes we have implemented (based on the results presented in this work) between Fall 2016 and 2017 affect student internalization of target material or the students' ability to develop strong insight?

Finally, we hope to answer the question of whether students are able to take knowledge of the design principles they obtain in AME4163 and apply them to novel challenges such as those they will encounter in their capstone design experience. Are they able to identify and test new principles of their own devising? How does their knowledge of design hold up in the absence of scaffolding in their project? We have collected data from a select group of students studied in Fall 2016 in a follow-up study over the course of their capstone design projects in Fall 2017 and are currently in the process of evaluating that information to address these questions.

6. Conclusions

We reiterate our central contention that the most important skill that will be required of the engineering graduate of 2020 will be the ability to adapt to the rapidly-shifting engineering landscape. Emerging new technologies, shifting paradigms, and increasing globalization will ensure that the most successful engineers are those capable of continuous learning and self-improvement. In our course, AME4163: Principles of Engineering Design, we seek to foster the ability to self-learn by implementing the LS to foster the identification of new knowledge through reflection on authentic, immersive experiences. Using LS data from the Fall 2016 iteration of this course, we demonstrate that individual students preferentially focus on design principles associated with team formation/management, concept generation, and critically analyzing the design process. Further, we confirm conclusions from our prior work that, on an assignment by assignment basis, students largely write about areas targeted by the assignment. Furthermore, we see that students whose LSs have been assessed by instructors as insightful tend to draw connections between the learning identified and specific future utility and are more likely to use words which demonstrate that connection. We are also satisfied that the LS instrument, employed in such a manner in an engineering DBT course such as AME4163, is useful both as a means of encouraging active reflection and for providing instructors with a window into the process by which they develop knowledge. Further, we improve on this approach with the integration of the text-mining method of data analysis, which has greatly improved both the turnaround time for analyzing LS data and the dimensions on which student learning can be assessed. We conclude by offering our approach and software used to engineering design educators interested in our work for educational or research purposes.

Acknowledgements—Jackson Autrey acknowledges the Graduate Teaching Assistantship from the School of Aerospace and Mechanical Engineering at the University of Oklahoma, Norman and the financial support granted through the Dolese Teaching Fellowship program. Jennifer Sieber acknowledges the Graduate Research Assistantship from the Systems Realization Laboratory at the University of Oklahoma. Farrokh Mistree acknowledges the financial support that he received from the LA Comp Chair. Zahed Siddique acknowledges the financial support that he received from the Dick and Shirley O'Shields Professorship in Mechanical Engineering.

References

1. J. L. Autrey, F. Mistree and Z. Siddique, Quantification of Learning Through Learning Statements and Text Mining in an Engineering Design, Build, and Test Course, *2017 ASEE Annual Conference and Exposition: Design in Engineering Education*, Columbus, OH, June 2017.
2. F. Mistree, Strategic Design Engineering: A Contemporary Paradigm for Engineering Design Education for the 21st Century? *Journal of Mechanical Design*, **135**(9), Sep. 2013, p. 090301.
3. J. Autrey, F. Mistree and Z. Siddique, Fostering Learning Principles of Engineering Design, *2016 ASEE Annual Conference and Exposition: Research on Design Learning*, New Orleans, LA, June 2016, Session M313.
4. D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development, Second Edition*, Pearson FT Press, Indianapolis, IN, 2014.

5. C. L. Dym, A. M. Agogino, O. Eris, D. D. Frey and L. J. Leifer, Engineering Design Thinking, Teaching, and Learning, *Journal of Engineering Education*, **94**(1), January 2005, pp. 103–120.
6. F. Mistree, J. Allen, H. Karandikar, J. A. Shupe and E. Bascaran, *Learning How to Design: A Minds-On, Hands-On, Decision-Based Approach*, 1995, https://www.researchgate.net/ publication/2818887_Learning_How_to_Design_A_Minds-On_Hands-On_Decision-Based_Approach accessed 5 April 2017.
7. L. Balmer, *Development of a Framework for Designing and Improving Courses that Embody Design-Specific Cognitive Activities*, M.S. Thesis, University of Oklahoma, Norman, OK, 2015.
8. Anon., *Commissions, A.E.A. Criteria for Accrediting Engineering Programs*, 2010, http://www.abet.org/uploadedFiles/ccreditation/Accreditation_Process/Accreditation_Documents/Archive/criteria-eac-2010-2011.pdf accessed 30 March 2017.
9. R. J. Eggert, Engineering Design Education: Surveys of Demand and Supply, *2003 ASEE Annual Conference and Exposition: Staying in Tune with Engineering Education*, Nashville, TN, June 2003, Session 3125.
10. B. Lahidji, Competencies in Manufacturing Engineering Technology Programs from Employer's Point of View, *2000 ASEE Annual Conference*, St. Louis, MO, June 2000, Session 3548.
11. T. J. Brumm, L. F. Hanneman and S. K. Mickelson, Assessing and Developing Program Outcomes through Workplace Competencies, *International Journal of Engineering Education*, **22**(1), 2006, p. 123.
12. M. A. Laingen, *Assessing Engineering Students' Demonstration of Workplace Competencies in Experiential Learning Environments through Internships and Cooperative Work Experiences*, Graduate Theses and Dissertations, Iowa State University Digital Repository, Paper 13957, 2014.
13. S. A. Male, M. B. Bush and E. S. Chapman, An Australian Study of Generic Competencies Required by Engineers, *European Journal of Engineering Education*, **36**(2), 2011, pp. 151–163.
14. E. De Graaff and W. Ravesteijn, Training Complete Engineers: Global Enterprise and Engineering Education, *European Journal of Engineering Education*, **26**(4), 2001, pp. 419–427.
15. R. H. Todd, C. D. Sorensen and S. P. Magleby, Designing a Senior Capstone Course to Satisfy Industrial Customers, *Journal of Engineering Education*, **82**(2), 1993, pp. 92–100.
16. H. A. Etlinger, A Retrospective on an Early Software Projects Course, *ACM SIGCSE Bulletin*, **22**(1), 1990, pp. 72–77.
17. W. F. Smith, Z. Siddique and F. Mistree, The Development of Competencies in a Design Course from a Student Perspective, *2014 ASEE Annual Conference & Exposition*, Indianapolis, IN, June 2014, Paper ID 9046.
18. M. Besterfield-Sacre, C. J. Atman and L. J. Shuman, Engineering Student Attitudes Assessment, *Journal of Engineering Education*, **87**(2), 1998, pp. 133–141.
19. M. Segers and F. Dochy, New Assessment Forms in Problem-based Learning: the Value-added of the Students' Perspective, *Studies in Higher Education*, **26**(3), 2001, pp. 327–343.
20. B. M. Olds, B. M. Moskal and R. L. Miller, Assessment in Engineering Education: Evolution, Approaches and Future Collaborations, *Journal of Engineering Education*, **94**(1), 2005, pp. 13–25.
21. J. Turns, Learning Essays and the Reflective Learner: Supporting Assessment in Engineering Design Education, *Frontiers in Education: 27th Annual Conference. Teaching and Learning in an Era of Change*, Pittsburgh, PA, November 1997, Session F1H, **2**, pp. 681–688.
22. J. K. Allen, F. Mistree, W. Newstetter and J. Turns, Learning Essays and the Reflective Learner: Supporting Reflection in Engineering Design Education, *1997 ASEE Annual Conference*, Milwaukee, WI, June 1997, Session 22300.
23. M. J. Frasciello, *Embedded Technical Communication Instruction: Understanding High-Context Writing in a Pre-Professional Engineering Program*, Dissertation, Syracuse University, New York, NY, 2015.
24. Y. F. B. Wu and X. Chen, eLearning Assessment through Textual Analysis of Class Discussions, *Fifth IEEE Conference on Advanced Learning Technologies*, Kaohsiung, Taiwan, July 2005, pp. 388–390.
25. B. J. Kokensparger, *Using Compositional Writing Samples to Explore Student Usage Patterns in a Learning Management System*, Dissertation, University of Nebraska, Lincoln, NE, 2013.
26. M. Choy, Effective Listings of Function Stop words for Twitter, *International Journal of Advanced Computer Science and Applications*, **3**(6), 2012, pp. 8–11.
27. S. M. Savaresi and D. Boley, On the performance of bisecting K-means and PDDP, *Proceedings of the First SIAM International Conference on Data Mining (ICDM-2001)*, Chicago, IL, April 2001, pp. 2–5.
28. M. Steinbach, G. Karypis and V. Kumar, A Comparison of Document Clustering Techniques[A], *KDD-2000 Workshop on Text Mining*, Boston, MA, August 2000, pp. 1–20.
29. I. T. Jolliffe, *Principal Component Analysis, Second Edition, Technical Report*, Springer-Verlag, New York, NY, 2002, pp. 1-6/10-27/78-85/111-115.

Jackson L. Autrey obtained his BS in Mechanical Engineering from the University of Oklahoma in 2014 and is currently an MS Student in Mechanical Engineering studying engineering education at the University of Oklahoma in the School of Aerospace and Mechanical Engineering. He is currently developing a doctoral dissertation proposal in engineering education and is a member of the American Society of Engineering Education. Mr. Autrey acknowledges the Graduate Teaching Assistantship from the School of Aerospace and Mechanical Engineering at the University of Oklahoma, Norman and the financial support granted through the Dolese Teaching Fellowship program.

Jennifer Sieber completed her BS in Microbiology at the University of Oklahoma in 2014 and obtained an MS in Data Science and Analytics from the University of Oklahoma in 2016. Her research interests include bioinformatics, specifically protein classification, neural networks for image classification, and text mining and classification of documents. She is currently employed as a data scientist focusing on metasearch optimization at Koddi Inc. in Fort Worth, Texas. Jennifer acknowledges the Research Assistantship from the School of Aerospace and Mechanical Engineering at the University of Oklahoma.

Zahed Siddique obtained his BS in Mechanical Engineering in 1994, his MS in Mechanical Engineering in 1996, and his PhD in Mechanical Engineering in 2000 from the Georgia Institute of Technology. His research interests include the development of tools and frameworks to enhance engineering design education, engineering design methodology, product platform design, virtual prototyping, and the use of computers in engineering design. In 2008, he won the Ralph R. Teetor Educational Award from the Society of Automotive Engineers and is a member of the American Society of Mechanical

Engineering, the Tau Beta Pi National Honor Society, the Golden Key Honor Society, and the International Association of Science and Technology for Development Technical Committee. Dr. Siddique currently holds the Dick and Shirley O'Shields Professorship in Mechanical Engineering, runs the Product and Process Design Lab, and is the Director of the School of Aerospace and Mechanical Engineering at the University of Oklahoma.

Farrokh Mistree obtained his Bachelor of Technology (honors) degree in Naval Architecture in 1967 from the Indian Institute of Technology and his PhD in Engineering from the University of California, Berkeley in 1974. He is the coauthor of two textbooks, one monograph, and more than 400 technical papers. Farrokh's research interests include the integrated realization of materials, products and associated manufacturing processes, the design of mechanical, thermal and structural systems; ships and aircraft; engineered networks; and engineering design education. He is a Fellow of the American Society of Mechanical Engineers, an Associate Fellow of the American Institute of Aeronautics and Astronautics, a Life Member of The Honor Society of Phi Kappa Phi, and a Member of the American Society of Engineering Education. Farrokh, received the ASME Ruth and Joel Spira Outstanding educator Award in 2011. Farrokh holds the L.A. Comp. Chair of Mechanical Engineering, co-directs the Systems Realization Laboratory, and is an adjunct professor in the School of Industrial and Systems Engineering at the University of Oklahoma.