ASEE's 123rd Annual • Conference & Exposition • New Orleans, LA • June 26-29, 2016

New Orleans JAZZED about Engineering Education

Paper ID #16406

Data-Driven Course Improvements: Using Artifact Analysis to Conquer ABET Criterion 4

Mr. Tony Andrew Lowe, Purdue University, West Lafayette

Tony Lowe is a PhD candidate in Engineering Education at Purdue University. He has a BSEE from Rose-Hulman Institute of Technology and a MSIT from Capella. He currently teaches as an adjunct at CTU Online and has been an on-and-off corporate educator and full time software engineer for twenty years.

David A. Evenhouse, Purdue University

David Evenhouse is a Graduate Student and Research Assistant in the Purdue School of Engineering Education. He graduated from Calvin College in the Spring of 2015 with a B.S.E. concentrating in Mechanical Engineering. Experiences during his undergraduate years included a semester in Spain, taking classes at the Universidad de Oviedo and the Escuela Politécnica de Ingenieria de Gijón, as well as multiple internships in Manufacturing and Quality Engineering. His current work primarily investigates the effects of select emergent pedagogies upon student and instructor performance and experience at the collegiate level. Other interests include engineering ethics, engineering philosophy, and the intersecting concerns of engineering industry and higher academia.

Mr. Dhinesh Balaji Radhakrishnan, Purdue University

Graduate Research Assistant at the School of Engineering Education, Purdue University

Dr. Stephen R. Hoffmann, Purdue University, West Lafayette

Stephen R. Hoffmann is the Assistant Head of the School of Engineering Education at Purdue University, with responsibilities for the First-Year Engineering Program.

Data Driven Course Improvements: Using Artifact Analysis to Conquer ABET Criterion 4

Abstract

This evidence based practice describes a process to evaluate a course within the spirit of ABET Criteria 4, continuous improvement. Faculty and staff are often asked to collaborate on the design and instruction of core engineering courses. Over time, these courses may evolve to accommodate new subject matter, pedagogical approaches, political and personal preferences, or other criteria as dictated by a dynamic group of stakeholders. Many changes originate from a clearly defined need or mandate, while others may sneak in without a full analysis of the course. Repeated and often subtle changes compound to have a significant impact on the course, creating a narrative reflecting the intents of the faculty and the concerns of the institution as course goals and methods are updated in each subsequent semester.

This paper describes a process to employ engineering education research methods to describe the nature, development, implications, and motivation behind of course changes. We define a six step process focused on the use of artifact analysis to provide instructional teams with concrete historical data, allowing them to better understand the structure of their course and how it has changed over time. A case study examining a large-format, First Year Engineering course is included at a part of this paper, providing context and serving to describe the process in action. The case study includes methodological choices, analysis, and findings as a guide to practitioners seeking to follow or further develop our process for gathering data. The data produced can be used to inform future changes to the course design to ensure alignment of the course objectives, assessment, and pedagogy, while at the same time systematically meeting the requirements of ABET Criteria 4.

Introduction

Many students begin their journey to becoming an Engineer in a classroom alongside dozens, if not hundreds, of their peers. These early courses are intended to present students with a set of core knowledge and skills that will prove useful across all engineering disciplines, thus molding the foundation of their academic careers. Year by year, thousands of students will go through this rite of passage in various class sections, with various instructors, eventually choosing between various engineering disciplines. It is not difficult to find examples of academic publications pertaining to the development, implementation, and performance of what we will call ‘large format courses’, a core course required for most if not all engineering students, taught by many instructors and likely designed and maintained by an instructional team[1–4]. In many of these cases, it could be argued that the most significant factors influencing student performance in, and perception of, a large format class come not from the students, nor the instructors, but from the curricular decisions of the cross-disciplinary course planning team charged with the design of the content, assessment and pedagogy employed within the classroom.

Literature concerning methods for high-quality engineering course design is well established, and should form the foundation of any initial course design, or major redesign, in an engineering curriculum. Experts advise that this practice begin with a thorough review of the learning objectives, subsequently taking into account assessment techniques to elicit learning and pedagogical methods that support both the content and objectives of the curriculum [5,6]. This approach is typically referred to as 'backwards design[7,8]. The goal of backward design is to first ensure each homework assignment, quiz, or test or other assessment is measuring how well the students demonstrate the desired learning objectives. Then the materials, media, activities, and other pedagogy are appraised to ensure they align to build the skills and understanding captured by the assessments. For typical first year engineering courses, the content of the course benefits from the experience and input of a wide array of interdisciplinary faculty. These faculty members essentially represent the course's constituency, and therefore should have a voice in the definition of learning objectives and other aspects of course development[9]. The same as when a practicing engineer completes problem scoping, an instructional team can take input from the diverse group of stakeholders, to define learning objectives. The backwards design process can then inform the rest of the course design for assessment and pedagogy.

As time passes, changes to curriculum always arise; areas for improvement will be spotted, materials will be modernized, new contributors will bring in fresh perspectives and energy, information or teaching methods will become obsolete. No course design is forever perfect, and it was perhaps this truism which led to ABET including "continuous improvement" as one of its criteria for engineering accreditation. Criterion 4 of the 2015-2016 ABET self-study questionnaire calls for institutions to "Describe how the results of evaluation processes for the student outcomes and any other available information have been systematically used as input in the continuous improvement of the program."[10]. When Criterion 4 was first revealed to be a part of ABET EC-2000 it proved to be quite concerning to college administrators tasked with maintaining accreditation. Continuous improvement's inclusion gave rise to a flurry of publications recommending how to satisfy ABET requirements, typically through the use of specific course design and course evaluation processes[5,9,11,12]. As the original ABET criterion called for institutions to employ assessment data for continuous improvement, early processes focused on student assessment data relating to the stated ABET learning objectives[5,9]. These recommendations were subsequently used as tools to reinforce the strength of a program's self-study for accreditation. However, some questioned whether focus on specific processes wasn't defeating the true purpose and spirit behind Criterion 4 and other ABET criteria[12].

Updating any course design based purely on a single data source like student assessment could easily prove contrary to the spirit of continuous improvement. Allow us to provide an example. For instance, say that a given cohort at an imaginary institution demonstrates particularly strong design skills during their sophomore year, but struggle in foundational math topics. Because of this, faculty in charge of the first year courses decide to update the first year engineering curriculum to address this issue. Just a couple years later, a new cohort comes to their sophomore year demonstrating exemplary math skills but poor design skills. By this time, a new group of faculty are in charge of the first year course, and they subsequently alter the first year curriculum to better emphasize design learning and deemphasize math skills. Such an infinite feedback loop would satisfy ABET's criteria by using assessments to inform changes in curriculum, but would hardly merit the label "continuous improvement".

Thus, a program that truly seeks continuous improvement should make changes which are backed by verifiable historical data. Likewise, instructional teams should weigh proposed revisions in light of the course's history. However, few course design or evaluation frameworks encourage instructors to review prior versions of the curriculum they are engaged with. For example, the popular book *The Systematic Design of Instruction* has little to no reference to reviewing documents or artifacts from past courses during the curriculum design[13]. Likewise, *Rethinking Engineering Education* makes reference to conducting a document review during course evaluation but overlooks a specific methods for conducting such a review [6]. More often, historical input is supplied as anecdotal evidence from experienced faculty, or less formal 'feedback'. For instance, a new instructor suggests adding in an active learning exercise to focus on student design, met with an off-handed, "oh, we tried that before" by the senior professor. Regardless of who is correct (it may be both parties are!), formal research could be employed to generate data and inform the course design process. Institutions often keep meticulous records of past courses, yet how often is this information archived and forgotten? Each term produces a bevy of documents in the form of syllabi, homework assignments, tests, and presentation materials, in addition to the student assessment data commonly studied by researchers. By revisiting these documents using artifact analysis, course designers can generate quantitative representations of the course. The resulting data provides justification for course revisions, and validity to programs claiming to pursue truly continuous course improvement.

This paper describes how to apply artifact review research methods to curriculum evaluation, thus providing historical data to course designers. After describing our suggested six-step process, we include a sample case-study evaluating a large-format, first year engineering course. This course is required for all incoming engineering students at Purdue University. The method presented draws from our experiences evaluating this course, as well as best practices described in literature concerning evaluation and course design.

Course Design Research Methods

A large format course combines the values of the university, the course designers, and the instructors. It must integrate the program outcomes of the departments it serves, the school's overall accreditation requirements, and its own course-specific learning objectives to create a cohesive whole. The identity and structure of large-format courses must be carefully crafted and accurately communicated. The course's faculty team will communicate their overall design to students and other stakeholders through the course description, syllabus, and other pedagogical materials. These documents become a static representation the course, and how it integrates with the broader curriculum and institution. While documents are always interpreted and mediated by each individual instructor in each classroom, they form a snapshot of the course in each offering. The method presented here looks to inform the work of course designers by providing a method to interrogate such course documents producing historical data.

Each artifact from previous terms acts as a source of data, allowing course designers to look for trends in design and execution. Together, multiple sources of data can indicate trends and identify areas in need of improvement. The data available within course documents may also be combined with data from traditional assessment strategies, such as student grades, in order to paint a more holistic picture of the course's execution. Feedback elicited from instructors or students through end-of-course surveys, interviews, or focus groups may contain tacit information to represents the implied understanding of the participants. Much like participant validation in more traditional thematic analysis methods[14], using feedback from the course's instructors can help to validate the results of an artifact review. It is important to recognize the variety of data sources available, from participants as well as artifacts, that can be used to improve the design of a course.

Artifact Analysis of Course Materials

Content analysis methods allow researchers to pull data from static documents. Saveyne and Robinson point out, "content analysis can be conducted on instructional materials: the lesson plans, classroom materials, assignments, and tests"[15]. Content analysis can be used to gather either quantitative or qualitative data, each paradigm using similar approaches and each potentially reinforcing the findings of the other. The general steps we would recommend for the informal analysis of course artifacts are:

1. Identify the objective of the analysis
2. Define of a coding scheme to captured desired data
3. Code the documents to elicit data
4. Verify coding reliability
5. Analyze data for trends
6. Validate the analysis

This method is generally standard to any content analysis, but here we are looking to describe its use for specifically understanding and informing curricular design.

Identify the objective of the analysis

The first step for the course designer is to identify the goal(s) of the analysis; is there already a clear curricular problem that needs addressing? Is the goal to evaluate the course as part of a continuous improvement process? The desired outcome for the analysis determines the type of data which is to be captured, a process known as coding. When coding, researchers carefully read the artifacts and search for the presence of specific words, themes, pictures or other contents. The coding scheme determines what the reader is seeking and how they will capture its presence (or lack thereof). When entering into the artifact review with a predefined problem in mind, the coding scheme will likely be *a priori*, a predefined series of data attributes the course designer is interested in capturing[16]. When the analysis is driven by a specific agenda it is easier to define the data desired to capture, and thus define a coding scheme. Alternatively, when the analysis is open to discovery or has no predefined agenda, the coding scheme can be *emergent*. Emergent coding allows the researcher to examining the documents and as they better understand the contents determine which attributes to capture for analysis[16].

Define of a coding scheme to captured desired data

Once the course designers determine their goals, choosing to employ *a priori* or *emergent* coding, the detailed coding scheme is established. For example, say instructors observe that older students are struggling to perform basic programing tasks, yet introductory programming was present as a learning objective for their first year course. The team would define coding attributes describing when and how often programming tasks were being assigned in the term, if these tasks are completed in teams or individually, or the weight assigned programming factors into the student's final grade over the years. These data points are considered valuable in analyzing trends year-by-year to see how instruction has changed. As the team begins reviewing documentations, they may decide to add in a count of, for example, how many programming samples are provided each year, further refining the data they deem valuable.

Code the documents to elicit data

Once the planning is completed, the research team's biggest task is to gather data. The team analyzes each artifact and applies the coding scheme to capture data. For basic documents this process can simply utilize paper notes or a spreadsheet. For more complex documents and coding schemes, specialized qualitative coding software can be used to coding documents and gather statistics. While the team will spend a majority of time reading and coding documents, this process can be spread across many individuals and can be shared amongst grad students, new faculty as well as established design team members to share the effort. The effort continues through all selected artifacts for each selected term being studied.

Verify coding reliability

Course designers should take time to check their data for reliability and validity. Reliability is a representation of how 'repeatable' the coding process is across different individuals as well as across time. If multiple researchers each code, or if a single individual codes, how consistently do they capture data between each other and from the first artifact to the last? Data reliability determines that data results can be trusted across the span of measures. Validity denotes how well the coded data represents what it was intended to measure. For instance, grades are only an accurate representation of what the students if the assessments are aligned to the outcomes. It is recommended that researchers employ more than one type of data to enhance validity [18], although such rigor is not possibly superfluous unless publication is a goal of the research. We will revisit the topic of validity later when we discuss combining the artifact analysis with informal interviews.

Reliability measurements ensure data is consistent across artifacts and particularly across years. The first, best method of ensuring reliability is to have multiple researchers code each document and compare results. This practice reduces bias by forcing discussion and consensus agreement, also leading to refined coding scheme. Two individuals may interpret documents distantly, and even a single coder may refine their interpretation over time. Reliability can start with discussion, but can also be calculated. Inter-rater reliability (IRR) is a statistical measure of how consistently two individuals coding results agree, while *intra*-rater reliability measures how consistently a single coder codes the same artifact over time. In formal research, a data set is set aside used to 'train' coders to produce stable and reproducible coding results. While it is beyond the scope of this paper for us to provide detailed instructions on how to determine IRR, Stemler[16] provides the exact formula for calculating IRR statistics, and a myriad of other resources and tools are available to help those wishing to verify coding reliability. Even for informal analysis however, IRR is a simple and valuable check that the data is telling a consistent story.

Analyze data for trends

The final stage of the content analysis is analyzing the data. The approach to analysis will vary depending on the coding scheme. Some attributes may be quantitative, lending to graphs or statistical analysis, some qualitative requiring other techniques such as thematic analysis[22]. An a priori coding scheme implies specific questions the course designers would like answered, informing the analysis. A multitude of resources are available to inform both qualitative and quantitative data analysis, and it is not within the scope of this paper to present them all here. A sample of what this analysis could look like can be taken from our case study detailed later.

Not all coded data needs to be discrete, ranked or ordinal. It is possible to look for topics or content within trends using summative content analysis[19]. Coders, or even a computer program, could be tasked to look for the presence of certain words or phrases and count their occurrence. Coding can even happen generally as a qualitative summary of, or quotation from, an artifact or portion thereof[18]. For such qualitative data, the course designer may wish to look for themes represented in the wording, rather than the trends we see in discrete data. If this proves difficult, the use of grounded theory[20] can help to elicit themes out of such attribute data.

Validate the analysis

The research data collected from an artifact analysis looks to obtain an objective perspective on course content, but does not necessarily provide an authentic look at the day-to-day practices in the classroom. The best source of relevant data on in-class instruction, beyond video evidence perhaps, is the instructional team itself. The course instructors can both validate coding scheme[16] and resulting data, as well as provide further insights to the researchers. Savenye[15] discusses the use of interviews to triangulate research findings, as well as describing practices for conducting quality interviews. We suggest first conducting content analysis on the data, then verifying observed trends using informal instructor interviews. As with every other aspect of such an artifact review, course designers must agree upon the level of rigor they think appropriate for their interviews. Hsieh[19] describes the use of recording and transcribing interviews for further qualitative analysis. Some designers may decide, as we did, that formal analysis of the interviews in overly rigorous for the purpose of informing course revisions. Formal interviews may be advisable if the artifact review is part of a broader research project. Most likely, the interviews can be semi-structured and informal, mixing predefined questions and free discussion of the data and its implications.

When applying our method described here, if the course designers, artifact reviewers, and instructional team are all the same people, interviews may seem superfluous. However, we would suggest that taking the time to document the attitudes and anecdotes of the instructors may help to generate new thoughts and insights on the data, regardless of their level of involvement in the actual analysis. The goal of our process is to generate data and observe trends in order to recommend improvements to the course, so it seems logical to socialize the data, discuss its implications, and plan actions accordingly. The interview, which may be simply the next scheduled meeting of the instructional design team, seems a clear step in validating the artifact review and may have the added benefits of inclusion and consensus building.

Applications of the Methods: A Case Study

To better describe our method, we hope a walkthrough of our case study may provide deeper insight to how artifact analysis might be executed, and the value it can provide to course designers. The object of study was the Fall semester, first year engineering course provided by Purdue University. Each sub-section below aligns to the six steps described in the method above, and describes how we carried out the step in our own analysis.

Identify the objective of the analysis

Our main source of data included materials and documents electronically archived between 2010 and 2015. These included each term’s syllabus, homework assignments, quizzes, and tests. Many pedagogical elements, such as presentation slides, were also archived, though we chose not to code these artifacts as they were employed at instructor’s discretion, giving us no guarantee that they were used identically across all of the multiple class sections. The instructor interviews confirmed that most instructors altered or declined to use these materials, preferring their own. Without a extrinsic mandate, we defined the objective of our research to be a general investigation into how well aligned the course assessments were to the stated course objectives, and secondarily, to elicit any insights we could on how the previous conversion to a flipped classroom format served to alter the course design over time.

Define of a coding scheme to captured desired data

Based on these objectives we reviewed the initial year’s materials, 2010, to identify coding attributes that were easy to acquire and prove valuable to our goals. Many of the data items we captured could be aptly described as ‘low hanging fruit’. In each homework assignment the document header clearly laid out expectations for who was to complete the assignment (an individual or team) and how was it to be submitted (paper or electronic). However, other attributes were noticeably more complex, requiring careful definitions to coding consistently. For instance, we wanted to capture what extent of homework assignments were simple plug-and-chug formula work, versus what extent of the assignments were comprised of general writing prompts, but also capture those assignments that existed somewhere in the middle, requiring students to write about technical engineering work. This requires definition and at least some training to code consistently. Table 1 details the attributes we coded in analysis of the homework assignments. Hopefully the concepts being pursued are clear through our struggles to find proper and acceptable names within each category.

Table 1: Homework assignment coding attributes

Assignment Attributes	Coding Categories
Individual vs. Team	Individual (1), Pairs (2), Team (3)
Hardcopy vs. Electronic submission	Hardcopy (1), Both (2), Electronic (3)
Technology Used	Excel (1), Both (2), Matlab (3)
Skills Required	Math/Science (1), Both (2), Communications (3)
Writing/Engineering Mix	General Writing (1) to Writing about Engineering (3) to Computational Only (5)
Bloom's Level of Learning	Remembering (1) to Creating (6) [17]
Alignment with Course Objectives	Not Aligned (1) to Perfectly Aligned (5)

Other coding attributes developed over time as we better understood the content and potential analysis. In fact, towards the end of our study we chose to go back and code for a new attribute, establishing which course learning objectives each homework assignment aligned to, as we initial only captured the overall alignment of each assignment to the course as a whole. Such emergent coding is additional work, but can prove valuable for the deep insight it can provide through reactive and iterative analysis. To avoid having multiple coding iterations, we strongly encourage thoughtful consideration of an initial coding schema. Researchers should consider the least satisfying aspects of the course and identify which aspects have been the most stagnant. Identify attributes which, when combined with data from student assessments, surveys, or other evaluation methods, can be used to paint a comprehensive picture of the course and its execution. The combination of planning with the methodological flexibility to evolve our process as needed provided us with a rich data set for analysis.

Code the documents to elicit data

Having validated the coding scheme for homework assignments within 2010 data, the process was repeated for each year's artifacts. Allowing for emergent coding categories allowed maturation of our data set. After coding and graphic two years of data, we recognized yearly-aggregate data alone did not tell the whole story. By adding the week the homework was assigned, we could present both aggregate view (Figure 1) and longitudinal analysis (Figure 2) of yearly data. Additional coding was conducted on the exams and syllabi for each year (not detailed here), but generally not valuable in analysis. Exams data was unusable as the format of the exams changed each year. The syllabi provided course objectives, coded using Bloom's taxonomy[17], but no additional relevant data.

Verify coding reliability

The reliability of these initial coding categories was verified using the 2010 artifacts by employing IRR. The reliability was then checked upon completion of coding for each subsequent year.

Analyze data for trends

Quantitative content analysis methods specialize in counting and analyzing things which, at first glance, may not seem quantifiable. In our case study, we captured a number of attributes that produced ordinal data. Thus, we were able to 'rank' assignments by their level of learning or alignment to course objectives, producing data which lends well to graphical representation. The first and simplest tool we employed for our analysis was Excel, which we used to generate graphical depictions of our data. This allowed for the quick identification of simple trends. Figure 1 provides an example, the percentage of assignments in the given year that aligned well (1-5) to the objectives. One interpretation of this data is a 'slipping' of alignment between homework and objectives from 2011 to 2014, followed by a correction in 2014. Triangulating this data with the changes to course objectives tells part of the story, and other sources can inform a fuller picture and perhaps inform desired actions. Further analysis using statistical software could lend further credibility to the data and any conclusions that might be drawn from them.

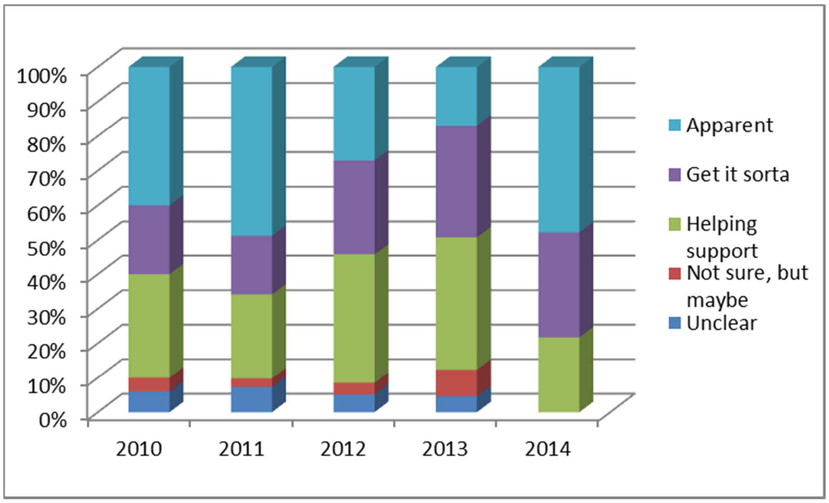

Figure 1: Yearly net perceived level of alignment between homework assignments and course objectives

In some cases, a finer view of the data may provide greater insight. To this end, we also graphically depicted data on a week-to-week basis using a custom software instrument. Figure 2 shows an example output from this tool, where the x-axis depicts weeks from 1-16, and the y-axis depicts the coded value. This chart depicts if students were asked to complete assignments as an individual (coded as a one), paired (coded as two), or team (coded as a three). This software was developed with no attention for aesthetics, yet provides quick views of the exact data desired. For example, a course designer looking at the image below would likely note that team assignments became much more common in Fall of 2014, with team tasks being assigned up to week 14. This may help determine how team tasks are distributed in the future to meet the course goals and outcomes.

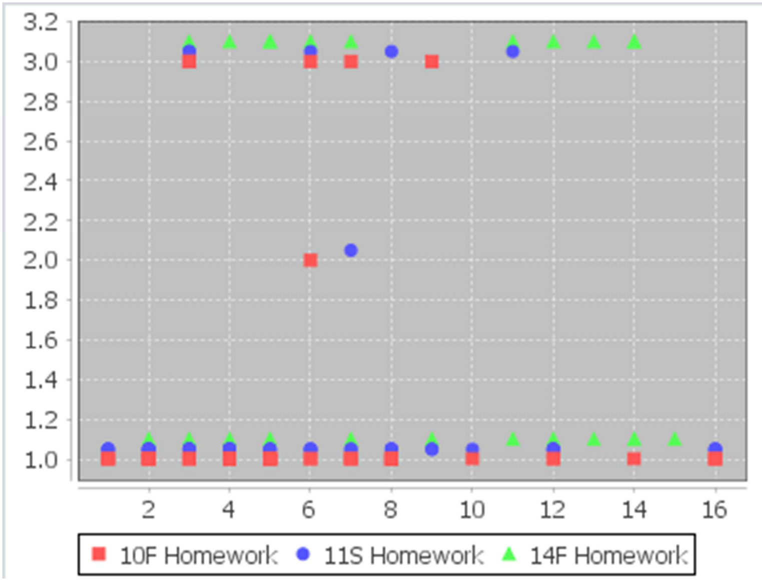

Figure 2: Weekly occurrence of individual, paired, and team assignments.

Using the graphs, we were able to make a number of observations about the changes in course artifact over time.

- Team assignments were more frequent, started earlier and continued later as the years progressed showing a greater emphasis on teamwork.
- Introducing computational tools (i.e. Matlab) seemed to disrupt the time spent on other learning goals, yet did not become a learning goal.
- As time went on, the course objectives reflected higher levels of learning (per Bloom), and the assignments generally changed to align with this.

Some of the graphs directly spell out trends. Figure 2 shows the later year (triangles) has more team assignments (3 on the y-axis) throughout the year than the others. Other interpretations need a combination of sources or analyses to tease out trends. For instance, Figure 3 compares course objectives within each of Blooms levels of learning alongside the same for homework assignments.

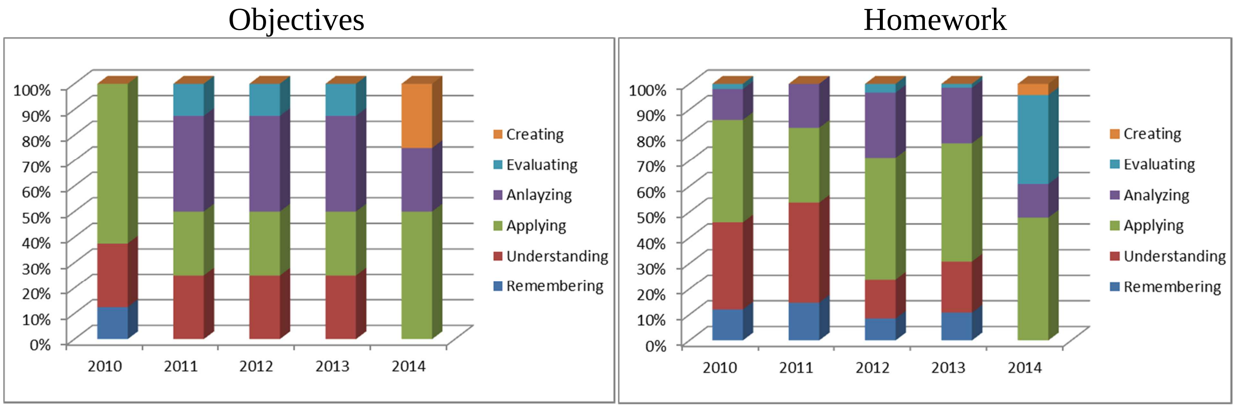

Figure 3: Bloom’s levels of learning evidenced in yearly assignments and course objectives

As the objectives move towards higher order goals, the homework gradually, but not entirely, follows suit. This finding could lead the course designers to investigate if the homework assignments are either providing tasks which scaffold students to these higher level assignments (confirming the design), or the course objectives are too ambitious and should be rewritten, or the homework require reworking to better challenge students and align to the stated learning objectives.

For one final example of a complex finding, we compared the alignment of the objectives and the use of technology. Within the course design we found a tension between the desire to have students performing more authentic design tasks (which are typically higher order learning objectives), and their need to be able to employ tools like Excel and Matlab. When computational tools were highly integrated, the number and quality of design tasks seemed to drop in number and alignment. As Matlab specifically was phased out of the course, the alignment returned. This observation is not to judge the use of Matlab, but gives valuable feedback to course designers on how well they may be achieving their stated course goals, or conversely, on how accurately they are depicting the content of their course through the course description and the syllabus. Fundamentally, the artifact analysis lends contextual data to the debate over proper course objectives and content.

Validate the analysis

In our study, we elicited feedback from former instructors in two ways. Given that we had little familiarity with the course, instructor feedback was vital to both our understanding of the data, and also the validation of our coding. Each instructor who agreed to speak with us was emailed a set of five predefined questions, the intent of which was to gather their perceptions on key aspects of the course before allowing them to see our data, which we feared may have biased their responses. After receiving their response, we conducted an informal interview to validate the content analysis methods, coded data, and longitudinal trends we found in our analysis. As the goal of the study was not a specific improvement, we did not suggest any changed. However, in the tradition of “life imitates art”, the instructional design team requested our data and a further meeting after we drafted this paper as to aid their course changes for next year!

Discussion

Even without an ABET criterion requiring institutions to be constantly improving their courses, change is inevitable. Engineers in industry look to make informed and intelligent changes to systems. They model options and use production data whenever possible to inform changes to product and process design. Engineering education research methods afford us tools to generate this same type of valuable data for those whose job it is to engineer the best possible educational experience. The content analysis methodology, as applied to artifact review, provides a way to reduce bias in decision making and measure the impact of change over time, rather than mistaking a reactionary change for inherent improvement.

This analysis method also holds the potential to elicit fresh, outsider viewpoints on the course content and materials. In our study, we did not receive any specific promptings or instructions after being granted access to the online course archive. We were forced to interpret the artifacts in light of our own, outsider perspective, later confirming or rectifying our observations with the help of instructor interviews. This experience could be analogous to that of a student, research assistant, or new instructor; removing the bias that may cloud the vision of experienced instructors or heavy stakeholders in the existing course. In many ways, conducting content analysis using graduate students or other outsiders can facilitate a boarder evaluation as well as remove the workload from the instructional team. Looking at the artifacts in their rawest form brings an outside perspective to the content, showing how the material is actually received by students and may elicit ideas not constrained by tacit departmental rules.

Our study did not capture data on student performance or perceptions, but such data can provide deeper insight and confirmation. Grades and course evaluations provide data on the student experience that can triangulate findings from the artifact analysis and further define areas in need of improvement. If the artifact analysis span many years, the changes in design can be compared to student performance in the class and possibly as the cohort progresses in their degree path. This would allow for better understanding of how course design changes impact overall student learning outcomes. We believe it possible to ground the redesign of courses in a wide variety of data, far beyond the rough methods proposed here. The combination of information on the course design, student perception, and student performance could grant the instructional team a valuable tool to inform design tradeoffs, drive change, and evaluate the results of learning interventions.

Conclusions

The design of any course will inevitably change over time. In foundational courses the subject matter may remain much the same, but other factors such as technology, politics, priorities, or prior student learning may change, forcing the pedagogy to adjust in response. The goal of the artifact analysis method explored here is to provide instructional teams with the tools they need to align their course design to the course objectives, and to generate discussion on the topic of data-driven course revision. We believe the best courses are designed to have a clear alignment between well-defined course objectives, properly constructed formative and summative assessments, and a contextually appropriate pedagogy to facilitate student learning.

The most important resource in the development of a course is its design team, but given limited time and resources it is important to make effective and well-informed decisions in a timely manner.

> Making a decision resembles climbing a mountain. The various paths to the summit are the alternatives. Biases and wishful thinking may lead to a disaster, and the professional climber will be careful to get the objective information about the paths, using aerial photographs, telescope observations, reports of previous climbing and weather forecasts. He will complement this information by his experience and other professionals' judgment, and will weigh the information with respect to his subjective evaluation of his fitness and capabilities, and according to his preferences regarding landscape, challenges and self-satisfaction.[21]

Engineering design is much the same; an iterative process seeking a unity of product form and function. It is often team-based, customer-driven, and utilizes feedback and performance data to continue designing improvements even after the product's initial release. It should be almost natural then, for the engineering educators of our time to seek continuous, data-driven improvement. We hope that the ideas, concepts, and the basic artifact review method detailed here can help facilitate such a mindset of continuous improvement, provide an approach to data gathering, and inspire others to begin developing their own tools and methods for enhancing student learning at the collegiate level.

Acknowledgements

We would like thank Purdue University School of Engineering Education for sponsoring our research as part of its Explorers program. We also would like to thank Richard Aleong, Chanel Beebe, Hoda Ehsan, Nathan Hicks, and Rose Xu for their feedback throughout the process and Dr. Michael Loui for expert guidance in social construction of knowledge.

References

1. Willson, V. L., Monogue, T. & Malave, C. First year comparative evaluation of the Texas A&M freshman integrated engineering program. in *1995 Frontiers in Education Conference* (1995).
2. Olds, B. M. & Miller, R. L. The effect of a first-year integrated engineering curriculum on graduation rates and student satisfaction: A longitudinal study. *Journal of Engineering Education* **93,** 23–35 (2004).
3. Ambrose, S. A. & Amon, C. H. Systematic design of a first-year mechancial engineering course at Carnegie Mellon University. *Journal of Engineering Education* 173–181 (1997).
4. Froyd, J. E. & Rogers, G. J. Evolution and evaluation of an integrated, first-year curriculum. *Proceedings Frontiers in Education 1997 27th Annual Conference. Teaching and Learning in an Era of Change* **2,** 1107–1113 (1997).
5. McGourty, J., Sebastian, C. & Swart, W. Developing a Comprehensive Assessment Program for Engineering Education*. *Journal of Engineering Education* 355–361 (1998). doi:10.1002/j.2168-9830.1998.tb00365.x
6. Crawley, E. F., Malmqvist, J., Östlund, S. & Brodeur, D. R. *Rethinking Engineering Education*. (2014). doi:10.1007/978-3-319-05561-9
7. Hansen, E. J. *Idea-based learning: A course design process to promote conceptual understanding*. (Stylus Publishing, LLC., 2012).
8. Wiggins, G. P. & McTighe, J. *Understanding by Design*. (ASCD, 2005).
9. Aldridge, M. D. & Benefield, L. A planning model for ABET engineering criteria 2000. in *1997 Frontiers in Education Conference* (1997). doi:10.1017/CBO9781107415324.004

10. ABET. ABET self-study questionaire: Template for a self-study report. (2014).
11. Felder, R. M. & Brent, R. Designing and Teaching Courses to Satisfy the ABET Engineering Criteria. *Journal of Engineering Education* **92,** 7–25 (2003).
12. Besterfield-Sacre, M. *et al.* Defining the outcomes: A framework for EC-2000. *IEEE Transactions on Education* **43,** 100–110 (2000).
13. Carey, L., Carey, J. & Dick, W. *The Systematic Design of Instruction*. (Pearson, 2005).
14. Long, T. & Johnson, M. Reliability and Validity in Qualitative Research. *Clinical Effectiveness in Nursing* **4,** 30–37 (2000).
15. Savenye, W. C. & Robinson, R. S. Using qualitative research methods in higher education. *Journal of computing in Higher education* **16,** 65–95 (2005).
16. Stemler, S. An overview of content analysis. *Practical Assessment, Research & Evaluation* **7,** (2001).
17. Bloom, B. S. & Examiners, C. of C. and U. *Taxonomy of educational objectives*. **1,** (David McKay New York, 1956).
18. Weber, R. P. Content Analysis. *Social Research Methods: A Reader* 117–124 (2004).
19. Hsieh, H.-F., Shannon, S. Three approaches to qualitative content analysis. *Qualitative Health Research* **15,** 1277–1288 (2005).
20. Birks, M. & Mills, J. Quality processes in grounded theory research. *Grounded Theory. A Practical Guide* 32–48 (2015). doi:10.1007/978-3-8349-9258-1_12
21. Buchanan, J. T., Henig, E. J., Aviv, T., Mordecai, I. & Henig, I. Objectivity and subjectivity in the decision making process.
22. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101.