ASEE AMERICAN SOCIETY FOR ENGINEERING EDUCATION

Paper ID #35706

The Use of Mixed Methods in Academic Program Evaluation

Mr. Michael B. O'Connor PE P.E., New York University

Michael O'Connor, Retired Professional Civil Engineer (Maryland and California), M.ASCE, is a member of the ASCE Committee on Developing Leaders, History and Heritage, Civil Engineering Body of Knowledge (CEBoK), and Engineering Grades. Michael has been a practicing Civil Engineer with over 50 years of engineering, construction, and project management experience split equally between the public and private sectors. Programs ranged from the San Francisco Bay Area Rapid Transit district's 1990's expansions in the East Bay and SFO Airport at three billion to the New Starts program for the Federal Transit Administration with over a hundred projects and $85 billion in construction value. At the latter, he also acted as source selection board chairman and program COTR for $200 plus million in task order contracts for engineering services. Working for the third-largest transit agency in the United States, the Los Angeles County MTA, Michael managed bus vehicle engineering for $1 billion in new acquisitions and post-delivery maintenance support for 2300 vehicles with some of the most complex technology (natural gas engines and embedded systems) in the US transit industry in the 1990s. Michael also has extensive experience as an instructor at New York University (five years), Howard University (four years), and California State University- San Francisco (ten years).

The Use of Mixed Methods in Academic Program Evaluation

Introduction

The Accreditation Board for Engineering and Technology (ABET) and Middle States Commission on Higher Education (MSCHE) are accreditation organizations. Together, their accreditation efforts assure the public that institutions and programs successfully prepare graduates to enter critical STEM fields in the global workforce.[1][2] Ongoing assessment and evaluation processes support engineering degree programs in higher education institutions.[3] This effort presents unique organizational and logistical challenges, such as integrating qualitative and quantitative data and supporting information in a meaningful way that demonstrates continuous improvement or the fundamental problem of integration. In 2020, the New York University, Tandon School of engineering, was readying for its ABET accreditation visit and its subsequent MSCHE visit in 2022-23. The preparation of the self-study report for ABET raised the aforementioned problem of integrating inputs and processes in a meaningful way that facilitates continuous improvement. The purpose of this paper is to describe how the New York University (NYU) BSCE program addressed ABET and MSCHE requirements and resolved the integration problem using mixed methods.

Accreditation, program control, and the fundamental problem of integration

ABET makes the argument for program control by explicitly stating that "institution(s) must demonstrate control over the (education degree) program to ensure compliance with all accreditation criteria and policies." [4] This paper defines the process of achieving that control over the institution's educational effectiveness as "program control." This means meeting ABET and MSCHE expectations on assessment, evaluation, and continuous improvement. Thus, an educational degree program such as the BSCE is controlled when it demonstrates its activities are compliant and its outcomes meet ABET and MSCHE expectations. [5] [6] Like ABET, MSCHE emphasizes a "culture of continuous improvement." [7] [8] The MSCHE guidance for accreditation consists of two parts- requirements for affiliation ('requirement") and standards for accreditation ("standard"). The MSCHE requirements and standards are in relevant part: [9][10]

- **Requirement 8** states that institutions must systematically evaluate their programs which means that they must document their objectives for student learning (attainment) and assess the level of attainment accomplished by their students.
- **Requirement 9** states that the coursework for the engineering degree program must be characterized by rigor, coherence, and appropriate assessment of the level of attainment accomplished by their students. This language is similar to Standard III, which adds more context in stating that programs must promote "synthesis of learning."
- **Requirement 10** states that the institution's evaluation process and continuous improvement efforts must be wholly integrated. This requirement means combining goals for student learning achievement with academic, institution, and third-party assessments. A relevant example of a third-party provider of assessment services for engineering degree programs is the National Council of Examiners for Engineering and Surveying (NCEES) and their Fundamental of Engineering (FE) assessment. Otherwise, this language is similar to Standard V.

- **Standard V** (Educational Effectiveness Assessment) requires institutions to deliver organized and systemic assessments (criterion 2) that evaluate student achievement of program goals and communicate those results to stakeholders. The institution is also required to demonstrate its integration of assessment results in its efforts to improve its educational effectiveness (criterion 3).

Hossain et al. (2019) argue that there is a

> "…fundamental difference between MSCHE (regional/institutional) and ABET (program-level) accreditations. MSCHE accreditation is an institutional process that grants credibility to the institution as a whole rather than to any specific program.
> …
> MSCHE accreditation requires an in-depth assessment of higher-level institutional activities, which are not needed (in-depth) by the ABET accreditation. In contrast, the ABET accreditation necessitates an in-depth assessment of technical competencies that are not required (in-depth) by the MSCHE accreditation." [11]

A clear sign of that emphasis on technical competency is that, unlike MSCHE, ABET standardizes the student learning outcomes for all engineering degree programs. (The SO1 thru SO7 for BSCE degree programs.) The statement about the in-depth nature of the assessment process misses the emphasis ABET places on evaluation and continuous improvement. MSCHE uses the word "continuous" fifteen times in its self-study report guidance. [12] Only two of those instances are for "continuous improvement," and they are only examples. Unlike ABET's criterion four, none of the MSCHE requirements or standards are titled "continuous improvement." For ABET, evaluation, in this instance, is being defined as a process (system) for interpreting assessment results [13] that are formalized (organized). It consists of making judgments based upon assessment data and information using approved faculty policies and criteria regarding program improvement. [14]

MSCHE and ABET differ in synthesizing assessment and evaluation data with programmatic decisions to maintain control over the engineering degree program. ABET says only that evaluation results and other available information must be systemically used as inputs for continuous improvement. [15] The institution must demonstrate changes or improvement plans based on this evaluation data.[16] MSCHE does not offer such guidance on continuous improvement but makes a more explicit requirement for synthesizing data and information, and continuous improvement efforts must be wholly integrated. This requirement means combining goals for student learning achievement with academic, institution, and third-party assessments. The key phrase in this statement is "completely integrated." Thus, program control can only be successful when the underlying inputs and processes are wholly and completely integrated in a meaningful way that facilitates continuous improvement. Therefore, the fundamental problem for engineering degree programs such as the New York University BSCE is defining and implementing an integration concept that enables program control and continuous improvement.

Program design and continuous improvement

One of the simplest ways to design a process to meet regulatory or accreditation purposes is to find examples of best practices, in this case, ABET self-study reports. [17] Several difficulties present themselves, and the first is the confidential nature of the process and data. ABET states that …

> “All information supplied (in the institutional self-study report) is for the confidential use of ABET and its authorized agents. It will not be disclosed without authorization of the institution concerned, except for summary data not identifiable to a specific institution or documents in the public domain.” [18]

Nonetheless, a limited number of institutions with engineering degree programs publish their self-study reports. The author conducted a web search in 2019-2020 and identified twenty engineering degree programs with publicly available information. (Note A) The material write-up in criterion four- continuous improvement would demonstrate how each program met ABET expectations. As one BSCE program put it:

> “There is a continuing effort to make assessments more standardized and quantitative … so that year-to-year trends can be identified and evaluated in a systematic manner.” (2015 ABET self-study report for BSCE program)

This same program noted that its portfolio of assessment measures contained a substantial amount of qualitative information, and its goal was to move these measures over to a quantitative basis. The report noted the difficulty the faculty was having in developing trend data with all the qualitative information. In the end, no detail was offered on how the program addressed the fundamental problem of integration discussed above.

The integration problem lies at the heart of the continuous improvement (CI) requirement. Azzouni et al. (2021) [19] conducted an extensive literature search on this topic, particularly in the academic accreditation setting. They note that CI can be analyzed and implemented in two ways, either as a system or a process with the ABET student outcomes criteria (SOs 1 thru 7) “…identify(ing) the major components...” CI data inputs can be either quantitative, qualitative, or both. Azzouni et al. (2021) identified four critical areas to consider in CI efforts; (1) CI components, (2) the scope of each component, (3) component data, and most importantly for this paper, (4) component integration. They closed their article by noting the

> “the contrast between the sheer number of case studies and the small number of empirical research studies in this topic area.”

While their paper focused on academic computing programs, this author argues that their analysis and findings are equally applicable to BSCE programs. The author conducted a more limited literature search than Azzouni et al. (2021) by focusing on “continuous improvement models“ on the American Society of Civil Engineering and American Society for Engineering Education websites. It revealed a similar lack of detail on component integration. One example is Kunberger et al. (2014) [20], describing the continuous improvement model at Florida Gulf Coast University (FGCU) for their engineering degree programs. They acknowledged the difficulty of the integration problem, stating that

> “… not all of these (assessment) methods factor evenly into the overall determination of student outcome achievement.” Kunberger et al. (2014)

The core of the problem for FGCU was combining quantitive results from direct assessments such as the NCEES FE exam and indirect qualitative measures such as senior exit surveys. The solution was to present the faculty with “…compiled data in each category and in a holistic manner so they can decide whether each category is above, at or below the standard.” [21]

Rahemi and Seth (2008) described a similar continuous improvement process model for aeronautical engineering used at Vaughn College of Aeronautics and Technology (VCAT). [22] In that paper, there was no discussion on component integration; the authors stated that

> "The program evaluator (department head or program coordinator) then determines what learning outcomes are not being adequately achieved and provides adequate action plans to improve them." Rahemi and Seth (2008)

Another approach was that used by Hussain et al. (2016) [23] [24], which identified the integration of direct and indirect assessment measures and other information as an essential element of continuous improvement. They advocate for an aggregate weighting factor scheme based upon Bloom's taxonomy. They noted that designing and implementing this model was not a trivial process and required extensive effort. In addition to published papers, ABET makes a limited number of self-study reports available at its annual symposiums in a "reading room." This material can be inspected but not copied.[25] At the 2019 symposium, attended by the author, three BSCE program reports were made available. Only one program described how it addressed the component integration problem. This self-study report integrated direct and indirect assessment measures to form a performance index. This index was the sum of (1) the average of all indirect measures and (2) the average of all direct measures multiplied by two. The program made an obvious decision to overweight the direct assessment measures versus the indirect. This weighted sum of indirect and direct measures was then divided by three to produce the performance index. The program control aspect of the process was identifying intervention points based upon pre-determined intervention criteria, such as the index dropping below a certain level. (2018 ABET self-study report for BSCE program)

It is important to note that ABET expects that any evaluation on individual student learning outcomes to be based in part upon at least one direct assessment measure. [26] ABET cites this as one of the main areas for non-compliance for engineering degree programs:

> "Overly heavy reliance on survey data or little direct evidence of outcome attainment in the methods used (programs should seek direct evidence when possible)." [27]

Danielson and Rogers (2007) note that this integration of direct and indirect assessment measures is significant. One of the key findings of this paper is that the fundamental problem of component integration discussed above is, in ABET terms, the complete integration of direct and indirect assessment data and other applicable information in a meaningful way that facilitates continuous improvement. For MSCHE, institutions must use "defensible standards" (Standard V) to evaluate student learning outcomes. Continuous improvement efforts must "reflect conclusions drawn from assessment results with clearly documented improvement processes" (Standard VI). The author argues that only CI efforts that produce results using thoroughly documented processes that are transparent and reproducible by third parties meet the MSCHE standards. Furthermore, meeting the ABET requirements for "systemic evaluation" and MSCHE requirements for "complete integration" requires the

> "…collection or analysis of both quantitative and/or qualitative data in a single study (self-study report) in which the data are collected concurrently or sequentially, are given a priority, and involve the integration of the data at one or more stages in the process.." [28]

Mixed methods and continuous improvement

For this paper, mixed methods as a methodology are defined as combining qualitative and quantitative data in some defined and systemic manner (Creswell & Plano Clark, 2011, as cited in Creamer, 2018). Mixed methods are more than the quantitative information and qualitative data in a single deliverable. Programs such as ABET or MSCHE require a comprehensive view of student learning. Mixed methods can weave the

> "… often disparate quantitative and qualitative perspectives together. … The inductive/deductive formulation refocuses the conversation from data types to analytical procedures." (Reeping, 2019)

Qualitative data can be coded quantitively, i.e., assigned indicative numerical values, typically using indices or scales. Quantitative data, on the other hand, consists of numbers. These values can then "…be manipulated to achieve greater insight into the meaning of the data and examine specific hypotheses." [29] The author argues that the integration requirement can be accomplished using mixed methods. Program evaluation can

> "…apply mixed methods approaches to evaluate a phenomenon of interest more holistically. By combining qualitative and quantitative inquiry, mixed methods program evaluation provides a mechanism for stakeholders to assess and evaluate programs systematically (Caracelli & Greene, 1997; Mertens, 2014)." [30]

Reeping (2019) gives the following examples that are useful for an academic program looking to meet ABET/MSCHE quality expectations:

- **Confirmatory** (triangulation)- A mixed-methods approach verifies an observation or hypothesis using multiple data sources concerning the same phenomenon.
- **Development**- Mixed methods leverage the results of one approach to inform the design of the following method (e.g., sampling, instrument development).
- **Evaluation/Intervention-** Multiple sources of evidence are collected for assessment and evaluation purposes (e.g., a course, program, activity, etc.; added by Creamer (2018) in her text).

Mixed methods have been recognized in medical research as a critical methodology for evaluating complex topics drawing on the strengths of quantitative and qualitative methods. Guetterman, 2015 noted that such integration could occur where quantitative and qualitative data are merged in a convergent manner that …

> "involves quantitative and qualitative data collection and analysis at similar times, followed by an integrated analysis." [31]

The purpose is to present the data …

> "in the form of a table or figure, a joint display, that simultaneously arrays the quantitative and quantitative results. A joint display is defined as a way to "integrate the data by bringing the data together through a visual means to draw out new insights beyond the information gained from the separate quantitative and qualitative results. (Guetterman, 2015)

Convergent designs merge qualitative and quantitative data from academic degree programs to understand complex phenomena better. Dickson (2011) states that integration focuses on **"**the concordance between qualitative and quantitative results." [32] Dickson argues that mixed methods can also validate ("triangulate" in ABET terminology) the integration of direct and assessment data.

A mixed-methods approach to integrating student learning goals and assessment data

ABET, in its guidance, calls for "triangulating" assessment measures to determine their validity and specifically using direct measures to validate indirect measures such as student or faculty course evaluations[33] which is a use of mixed methods. MSCHE requires integrating assessment data (quantitative and qualitative) with attainment goals/objectives, again a mixed method. Mixed methods/convergent design was used to achieve the required component integration. The research found several approaches could be used to merge the qualitative and quantitative components.

- Wholistically review the results and make a summative judgment on the degree to which student attainment of learning had been accomplished. This was not recommended for the BSCE program as it was judged not to meet MSCHE requirements for "defensible standards" (Standard V) and documented improvement processes" (Standard VI). This is also a common issue in ABET reviews as an assessment/evaluation process that is "ad hoc" or inconsistently used.
- Create scaling for each component and then aggregate these into a single performance management index number, whether equally or unequally weighted. The program control aspect of the process was identifying intervention points based upon pre-determined intervention criteria, such as the index dropping below a certain level. This was also not recommended as it was judged as not meeting ABET requirements to an adequate level of disaggregated data for program evaluation and continuous improvement. The attainment goals or performance baselines were only set at the summary level. More extensive interventions into the program to cure deviations from baseline performance couldn't be supported.
- Triangulation design seeks to "… bring together the differing strengths and non-overlapping weaknesses of quantitative methods (large sample size, trends, generalization) with those of qualitative methods (small N, details, in-depth) and directly compare and contrast quantitative statistical results with qualitative findings or to validate or expand quantitative results with qualitative data." [34] This approach was recommended for the New York University BSCE program.

Using the triangulation design method, the integration problem for NYU's BSCE program requires aligning the data sets on "common themes" (Guetterman, 2015). The ABET student outcomes (SO1 thru SO7) are a suitable framework for accomplishing this goal. The quantitative and qualitative data can be jointly displayed and aligned on a common student outcome basis. In the 2020 NYU ABET SSR, the data table was vertically arrayed with longitudinally displayed data for the previous six years. (See Table 1 below for a sample.) The immediate problem then becomes constructing the array as the scaling for the data sets is different. One of the data sets is the test results from the Fundamentals of Engineering (FE) exam administered by NCEES. The Department chose to adopt the FE and the institutional comparator ratio score as the base for measuring attainment.[35] The learning goal adopted by the BSCE program was to exceed the mean performance of all other institutions with students taking the FE, a comparator score of greater than 1.0. This choice meant that the attainment goal for each particular FE subject area would also be greater than 1.0. This assessment component was integrated with the other components: student course evaluations (SCE) and Senior exit interviews (SEI). Both were surveys that posed questions to students about their self-reported assessment of learning attained on an outcome basis. The SCE and SEI instruments used a five-point Likert scale, and the Department chose to make the goal of 4.0 for both measures. For all three measures, if the scoring did not meet expectations, there was a requirement that the slope of the trend line fit to the data points by linear regression should be positive and greater than or equal to 0.010. (Note B)

The FE data sets could be scaled to a common framework such as the FE comparator ratio and report a single index number. For the SCE and SEI datasets, the students self-report their sense of agreement/disagreement with a statement ("I am now capable of solving complex problems"). For academic program purposes, the fundamental integration problem is how to mix external performance data such as the FE with an intensity scale of belief or range of responses to a survey ("Likert-type scale") and other available information such as faculty course-level assessment artifacts such as FCARs.[36]

The NYU BSCE program adopted a blended variable strategy for mixing the quantitative and qualitative program data. Reeping (2019) defines this as a type of transformation that consolidates qualitative and quantitative data into one new variable or category. The new variable, in this case, was "ranking" the data sets. For the FE data, each subject area comparator index was summed by ABET SO and ranked highest to lowest. For the SCE and SEI data, each measure assigned responses to specific ABET student outcomes (SOs). This was averaged for each SO and then ranked. The result is an n-tuple sequence of data elements, in this case, a 3-tuple- (FE ranking, SCE ranking, and SEI ranking).

MSCHE requires a defensible logic to mix these n-component rankings. The FE exam is well documented as a direct assessment measure and therefore needs no defense, but what about mixing FE data with the SEI and SCE survey measures? What is the link between direct assessment performance and student surveys? The role played by the student forming an identity of themself as competent professionals promote learning and performance. [37] The primary means of measuring such identity development and related performance are indirect assessment measures, primarily surveys. [38] Patrick et al. (2018) argued that a framework for identity was composed of the student's self-interest, recognition by others as a good engineering student, and confidence in their performance/competence to solve problems and understand engineering content essential to the development of engineering identity. The authors also found a positive correlation between identity and engineering performance. For this paper, the causal logic of mixing the 3-tuple- (FE ranking, SEI ranking, and SCE ranking) is that they share a common performance element and goal. The FE performance goal is to outperform the mean of its institutional peers. The performance goal of the Likert scale responses in the SCE and SEI measures is to promote an average of all responses greater than 4 (Agree). Creating one n-tuple for each ABET SO will provide the program with evidence of third-party assessment of student competence against their self-perception of the same.

Given the positive correlation established by search in this area, ideally, each n-tuple sequence of data elements, in this case, a 3-tuple- (FE ranking, SEI ranking, and SCE ranking) the rankings would be equal. That is, an individual SO would have all three measures with the same rankings (**1,1,1**) [36], i.e., "interpretative consistency." SOs with different rankings give rise to interpretative challenges when student attainment in third-party assessment and student selfreporting does not align and are not consistent. The issue of consistency is separate from non-attainment. The FE subject areas allocated to an individual ABET SO may not meet the program objective of outperforming the median value of other institutions irrespective of student perception. Alternately, FE performance may meet attainment, but student perception of self-performance (their engineering identity) does not align in the rankings. Significant in this interpretation process is (i) which measure ranks higher and (ii) the rankings' distribution or spread. The question for the NYU BSCE program is how much misalignment of ranking within the n-tuple can be tolerated without requiring intervention to correct the problem. This was resolved using the criteria and decision rules discussed below.

The use of mixed-methods for continuous improvement in the NYU BSCE program

The NYU BSCE program implemented a mixed-methods approach to evaluating its three main assessment measures for its 2020 ABET Self-study report (SSR) and continuous improvement inputs. Mixed-methods analysis and evaluation were performed for all seven ABET student outcomes. Four outcomes had three components (SO1-Complex problems, SO2-Design, SO4-Professional and ethical responsibilities, and SO6-Experimentation). An idealized example of a highly consistent alignment of the measures would be where all three meet their individual attainment goals and have the same ranking. This would be symbolized as a **1-1-1** "vector" in the form of an FE-SCE-SEI tuple. All three assessment measures (FE, SCE, and SEI) are expected to meet the specific goals for individual year scores as well as for six-year averages (FE ≥ 0.99, SCE ≥ 3.99, and SEI ≥ 3.99)
In the case of the 2020 NYU ABET SSR, some exemplary data will demonstrate the process.

	AY2014 (-Spring 2014)	AY2015 (Fall 2014-Spring 2015)	AY2016 (Fall 2015-Spring 2016)	AY2017 (Fall 2016-Spring 2017)	AY2018 (Fall 2017-Spring 2018)	AY2019 (Fall 2018-Spring 2019)	AY2020 (Fall 2019-)	Average	Ranking highest to lowest average (May 2020)	Slope	Intercept	RSQuared
[redacted]												
C3-SO1-FE	1.012	0.964	1.020	1.037	0.997	0.996	1.064	1.013	FE2	0.007	0.986	0.008
C3-SO1-SCE-A33-41			4.230	4.220	4.370	4.000	4.440	4.252	SCE1	0.020	4.172	0.035
C3-SO1-SEI		4.245	4.105	3.956	4.193	4.286	4.133	4.153	SEI2	0.006	4.089	0.010
C3-SO2-FE	1.050	0.941	1.022	1.015	0.986	1.028	1.124	1.024	FE1	0.013	0.954	0.184
C3-SO2-SCE-A33-41			4.170	4.020	4.330	3.780	4.570	4.174	SCE3	0.056	3.950	0.087
C3-SO2-SEI		3.930	3.753	3.615	3.908	3.643	4.067	3.819	SEI7	0.018	3.754	0.038

Table 1 – Sample evaluation data from criterion four analysis in the 2020 NYU ABET SSR

Table 1 shows SSR data for ABET student outcomes 1 and 2, all a part of criterion three. The code "C3-SO1-FE" as an example means ABET criterion three, student outcome one, FE exam performance. Table 1 depicts the longitudinal performance of AY2014 thru AY2019 for the three components. “C3-SO1-SCE-…” is for the course evaluations, and likewise, “C3-SO1-SEI” is for the senior exit interviews. The pattern is repeated for ABET student outcome 2 with “C3-SO2-FE”, etc. For SO1, all three components met their individual performance goals as a yearly average, although there is a potential concern over the slope of the performance data. For SO2, two out of three measures met their goals, but a third (SEI) did not. However, the SEI data does meet the program goal of an increasing slope, giving confidence to the assessment by the program that the measure will eventually come back into compliance.

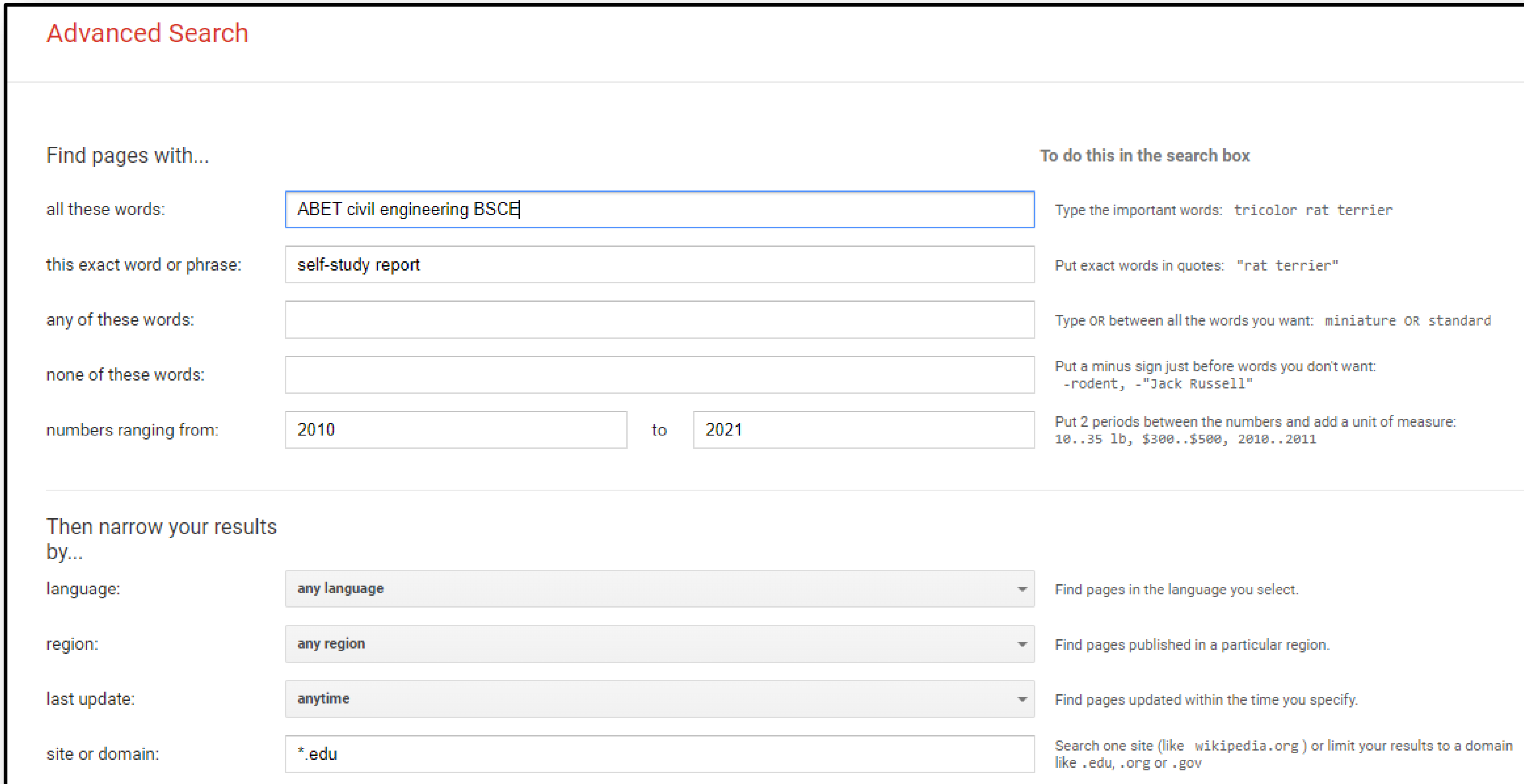
From an interpretative consistency point of view, the FE and SEI measures for ABET SO1 are ranked second highest in student outcome attainment, with the SCE having the highest ranking. The actual vector, in this case, was **2-1-2**. Since all three assessment tools meet the individual goals and the rankings are very high, the program argued in the SSR that there is a strong consistency among the FE, SCE, and SEI results. Therefore, student attainment objectives had been met, and the ABET program evaluators agreed in 2021 with this finding.
A different example of interpretative consistency was for ABET SO2 with an actual vector of **1-3-7**. The FE and SCE assessment measures meet the specific goals for most individual scores and six-year averages, but the SEI six-year average of 3.819 did not. The slope of the SEI trend line was positive and greater than 0.01, although the data showed substantial scatter. The FE exam ranks first out of four (FE1), the SCE ranks third out of seven (SCE3), and the SEI ranks lowest out of seven (SEI7). This data shows a persistent inconsistency between the FE exam and SCE data and the SEI data. The NYU BSCE program argued in its SSR that while the SO1 attainment was highly aligned, the SO2 misalignment required intervention to correct the problem. Reviewing several years of the SEI documentation has led to the recommendation that more context be provided to the Seniors to inform their SEI answers better and follow-up interviews. Similar context issues may also explain divergences in the SCE responses and allow students to better inform their answers in such surveys.

Conclusions and future work

The significant finding of this paper is that this approach to academic program evaluation by NYU is innovative and constitutes a new application of mixed methods relevant to the engineering education community. It also presents recommendations for applying the mixed-method approach in an environment where programs that conduct assessments must meet ABET and MSCHE requirements for continuous improvement. The NYU BSCE program has used mixed methods results in its evaluation processes resulting in decisions and actions for academic improvement. BSCE programs are often confronted with the problem of mixing external performance data with multiple data sources such as student surveys and interviews and Faculty analysis products (CARs). The n-component rankings approach detailed in this paper has a defensible logic in common performance elements and goals. Lastly, the n-tuple model offers a clear pathway to easily integrate more components into the evaluation process as programs develop more assessment components and establish additional decision rules to address the requirements of interpretive consistency.

Notes

A- The web search was conducted using advanced search query techniques. A copy of the search settings and results are shown below.



This query design resulted in 186 “hits,” and approximately twenty were the actual self-study reports. Other results related to ancillary documentation, resumes, announcements, etc.

B – The NYU BSCE SSR, in most cases, used AY2014-2020 data. In the case of the NCEES FE data, this had partial data sets, and AY14 only had Spring 2014 and AY20 only Spring 2020. After the SSR was issued in 2020, the faculty determined that limiting the longitudinal data analysis to the last six years was appropriate.

Acknowledgments.
The NYU Tandon School Civil and Urban Engineering Department Chair, Dr. Magued Iskander, provided comments and suggestions on the paper title and permission to use the 2020 BSCE program data. Lawrence Chiarelli and Edwing Medina provided comments and suggestions. Lastly, many faculty members and students collected the data that made these reports possible. Opinions, findings, and conclusions in this article do not necessarily reflect the views of New York University.

References

[1] [Online]. https://www.abet.org/accreditation/what-is-accreditation/why-abet-accreditation-matters/ [Accessed in 2021.]

[2] Hossain, Hossain, and Kouar, "Optimizing assessment tasks for institutional and program-level accreditations: A case study of accreditation requirements of msche and abet," *Journal of Assessment and Institutional Effectiveness*, vol. 9, no. 1-2, pp. 96–120, 2019.

[3] A. Hussain, T. Nizar, J. Nayfeh, and S. El-Nakla. "Undergraduate Engineering Program Evaluation, Assessment, and Continuous Improvement Process: A Case Study." In 2020 Gulf Southwest Section Conference. 2020.

[4] Accreditation Policy and Procedure Manual (APPM), ABET, 2020-2021

[5] [Online]. Accreditation Policy and Procedure Manual (APPM), 2020-2021, I.C.1.b- Eligibility of Programs for Accreditation Review. https://www.abet.org/accreditation/accreditation-criteria/accreditation-policy-and-procedure-manual-appm-2020-2021/ [Accessed in 2021.]

[6] [Online]. MSCHE Self-Study guide, The Commission's Expectations for Assessment, page 17 of 90. (2018) https://www.cmu.edu/middlestates/documents/selfstudyguideversion13a.pdf [Accessed in 2021.]

[7] [Online]. MSCHE Standards for Accreditation and Requirements of Affiliation, page 5 https://www.msche.org/standards/ [Accessed in 2021.]

[8] MSCHE Self-Study guide, The Commission's Expectations for Assessment, page 17 of 90. (2018)

[9] MSCHE Self-Study guide, Table 1: Requirements Of Affiliation, page 11 of 90. (2018)

[10] https://msche.box.com/shared/static/6upfla8coxha663p0j10u3gatow38jel.pdf

[11] Hossain, Hossain, and Kouar, page 104, (2019)

[12] MSCHE Self-Study guide, The Commission's Expectations for Assessment, page 17 of 90. (2018)

[13] Anon., "Self-study report development workshop" participant workbook, presented at the 2019 ABET symposium, 2019, page 4 of 64.

[14] ABET, 2019, page 4 of 64.

[15] ABET, 2019, page 22 of 64.

[16] M. B. OConnor, L. Chiarelli, "Faculty course assessment report- its evolution and future." 2021, unpublished.

[17] A. Tarantino, "Governance, risk, and compliance handbook: technology, finance, environmental, and international guidance and best practices". John Wiley & Sons, 2008.

[18] [Online] Template For a Readiness Review Report for the 2022-2023 Readiness Review Cycle, Extracted from 2021-22 EAC Self-Study Questionnaire, ABET Engineering Accreditation Commission (EAC) https://www.abet.org/wp-content/uploads/2021/09/E005-EAC-Self-Study-Questionaire-for-Masters-Level-2022-2023-cycle.docx [Accessed in 2021.]

[19] A. Azzouni, Abdullah, and J. Parham-Mocello. "Continuous Improvement in Academic Computing Programs is Rarely Comprehensive." *In 2021 ASEE Virtual Annual Conference Content Access*. 2021.

[20] T. Kunberger, K. Csavina, and R. O'Neill. "Scaffolding Assessment from Lesson Objectives through Student Outcomes: Faculty Perspectives of a Viable Continuous Improvement Model." *Journal of Professional Issues in Engineering Education and Practice* 140, no. 4 (2014): A2513001.
[21] T. Kunberger, K. Csavina, and R. O'Neill. Pp. 5, (2014)
[22] H. Rahemi, and N. Seth. "Student Learning Outcomes: An Integrated Continuous Improvement Process for Course and Program Assessment." *Latin American & Caribbean Journal of Engineering Education* 2, no. 2 (2008).
[23] W. Hussain, F. K. Mak, and M. F. Addas. "Engineering program evaluations based on automated measurement of performance indicators data classified into cognitive, affective, and psychomotor learning domains of the revised bloom's taxonomy." In *2016 ASEE Annual Conference & Exposition*. 2016.
[24] W. Hussain, M. F. Addas, and F. Mak. "Quality improvement with automated engineering program evaluations using performance indicators based on Bloom's 3 domains." In *2016 IEEE Frontiers in Education Conference (FIE)*, pp. 1-9. IEEE, 2016.
[25] [Online] 2015 ABET symposium program book, Self-Study Report Room, page 22 of 100, https://www.abet.org/wp-content/uploads/2015/04/2015-ABET-Symposium-Program-Book.compressed-1.pdf [Accessed in 2021.]
[26] S. Danielson, and B. Rogers. "A methodology for direct assessment of student attainment of program outcomes." *In 2007 ASEE Annual Conference & Exposition*, pp. 12-63. 2007.
[27] Anon., "Program Assessment workshop" participant workbook, presented at the 2019 ABET symposium, 2019, page 83 of 112.
[28] M. Borrego, E. Maura, P. Douglas, and C. Amelink. "Quantitative, qualitative, and mixed research methods in engineering education." *Journal of Engineering education* 98, no. 1 (2009): 53-66.
https://www.cs.auckland.ac.nz/courses/compsci747s2c/lectures/borrego.ea-quantitative-jee-09.pdf [Accessed in 2021.]
Citing Creswell, J.W., V.L. Plano Clark, M.L. Gutmann, and W.E. Hanson. 2003. *Advanced mixed methods research designs*. In Handbook of mixed methods in social and behavioral research, eds. A. Tashakkori and C. Teddlie, 209-240. Thousand Oaks, CA: Sage Publications.
[29] Qualitative vs. Quantitative, Trochim, Loyola Marymount University, Accessed on August 13, 2021
[30]] D. Reeping, A. Taylor, D. Knight, and C. Edwards. "Mixed methods analysis strategies in program evaluation beyond "a little quant here, a little qual there." *Journal of Engineering Education* 108, no. 2 (2019): 178-196., citing (Greene, Caracelli, & Graham, 1989))
[31] T. Guetterman, C. Fetters, M. D., & J. W. Creswell, "*Integrating Quantitative and Qualitative Results in Health Science Mixed Methods Research Through Joint Displays*." Annals of family medicine, 13(6), 554–561. https://doi.org/10.1370/afm.1865 (2015).
[32] V. V. Dickson, C. Lee, B. Riegel, *"How do cognitive function and knowledge affect heart failure self-care?"* J Mixed Methods Res.2011;5(2):167-189

[33] Anon., “Program Assessment workshop” participant workbook, presented at the 2019 ABET symposium, 2019, page 66 of 112.
[34] J. W. Creswell, and V. L. Plano Clark. “Designing and conducting mixed methods research.” Sage publications, 2007. Chapter 4
[35] [Online] B. Crawford, J. Steadman, D. Whitman, R. Young, *“Using the Fundamentals of Engineering (FE) Examination as an Outcomes Assessment Tool”, National Council of Examiners for Engineering and Surveying”* National Council of Examiners for Engineering and Surveying, https://ncees.org/wp-content/uploads/FEOAT-white-paper-2021_final-1.pdf [Accessed on October 4, 2021]
[36] Estell, John K. "The faculty course assessment report." In 33rd Annual Frontiers in Education, 2003. FIE 2003., vol. 1, pp. T4B-T48. IEEE, 2003
[37] A.Kaplan, and H. Flum. *"Identity formation in educational settings: A critical focus for education in the 21st century."* Contemporary Educational Psychology 37, no. 3 (2012): 171-175.
[38] A. D., Patrick, and A. N. Prybutok. *"Predicting persistence in engineering through an engineering identity scale."* International journal of engineering education 34, no. 2a (2018).