ABET Criterion 3 as a Measure of Engineering Self-Efficacy:

Comparing the New Criteria (3.1-7) to the Previous Criteria (3a-k)

Mark Schar
Stanford University
Stanford, CA USA
mfschar@stanford.edu

Patsy Brackin
Rose-Hulman Institute of Technology
Terre Haute, IN USA
brackin@rose-hulman.edu

K. J. Chew
Stanford University
Stanford, CA USA
kaijunch@stanford.edu

Sheri Sheppard
Stanford University
Stanford, CA USA
sheppard@stanford.edu

Abstract — **This Research to Practice Full Paper outlines the process of converting the ABET Criterion 3 - Student Outcomes into an engineering self-efficacy scale (ESE-ABET) and in this format compares the current ABET 3.1-7 items with the previous ABET 3a-k items with undergraduate engineering students. In total, survey responses were collected from undergraduate engineering students (n = 561) attending five institutions ranging from larger public universities to smaller private universities. Mean comparison, correlation and regression suggests that the ABET Criterion 3 do make valid and reliable measure of overall "engineeringness" self-efficacy and that self-efficacy scales based on ABET 3.1-7 and ABET 3a-k produce nearly identical results. Factor analysis and hierarchical clustering show that the ABET Criteria 3 items tend to form two groups - "Technical Skills" and "Professional Skills" - and this finding may make it easier to use the criteria as an organizing curricular structure and/or a measure of program effectiveness. Support was also found for a short form version (ESE-SF) of this scale when survey capacity is limited. A discussion of the value of having a valid and reliable ABET-based self-efficacy measure is included.**

Keywords — self-efficacy; engineering self-efficacy measures; ABET Criterion 3 - Student Outcomes; hierarchical clustering; curriculum development; ABET accreditation

Engineering educators have as their highest priority the growth and development of their students into fully capable engineers who can drive innovation and thrive in today's complex, technical world [1]. The measurement of student progress against this goal is challenging. Engineering education is complicated, spanning over 40 different academic majors with students of differing tenure and accomplishment [2].

This paper describes the development of a psychometric scale based on ABET Criteria 3 - Student Outcomes that can be used to measure perceived engineering capability or "engineering self-efficacy" for any engineering student at any stage of their development. This also compares previous ABET Criteria (ABET 3a-k) to newly established Criteria (ABET 3.1-7) to determine the extent that the new criteria cover the same skills areas as the previous criteria.

I. ABET General Criterion 3 – Student Outcomes

A. Background

ABET (formerly known as the Accreditation Board for Engineering and Technology) was established in 1932 as a non-governmental organization that accredits post-secondary education programs in "applied science, computing, engineering, and engineering technology" [3]. In 2017, ABET reports accrediting over 3,700 programs at over 750 colleges and universities in 30 countries [4]. Input to ABET accreditation criteria is broad based with contributions from over 2,200 volunteer academic and industry professionals from 35-member societies.

ABET assigns accreditation to four commissions tailored to the nature of the academic program. Each commission defines a set of eight criterion for accreditation ranging from categories like Students, Student Outcomes and Curriculum, to Faculty, Facilities and Institutional Support [5]. For the purposes of this study we are focusing on Criterion 3 – Student Outcomes for the Engineering Accreditation Commission (EAC).

B. ABET 3.1-7 and ABET 3a-k

Criterion 3 – Student Outcomes was substantially revised in 1996 to focus on eleven engineering student behaviors (known as ABET 3a-k) that were deemed important to engineering success in industry post-graduation [6]. The goal of establishing Criterion 3a-k was to define student behavioral outcomes that would be "common criteria for all engineering programs," [7] and the broad and inclusive nature of the ABET definition process helps insure that these criteria, in their totality, are a relevant definition of "engineeringness" for all accredited programs.

While these criteria have inspired positive changes to student outcomes [8] they have also been subject to continual improvement and adjustment. Starting in 2009, the EAC Commission has now completed a "second generation" revision process for the Criterion 3 behaviors (now referred to as ABET 3.1-7) resulting in seven (versus eleven) revised criteria [9]. For the purposes of this study we compared the previous ABET 3a-k to the new ABET 3.1-7. The complete criteria statements of both ABET 3a-k and 3.1-7 are shown in TABLE I.

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

TABLE I. ABET Criterion 3: Previous 3a-k and New 3.1-7 Criteria

GENERAL CRITERION 3. STUDENT OUTCOMES:
The program must have documented student outcomes that prepare graduates to attain the program educational objectives.

Code	Label	Description
Current Criteria		
3.1	solve complex	identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics
3.2	engineering design	apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors
3.3	communicate	communicate effectively with a range of audiences
3.4	ethical solutions	recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts
3.5	team leadership	function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives
3.6	experiment analyze	develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions
3.7	learning strategy	acquire and apply new knowledge as needed, using appropriate learning strategies
Previous Criteria		
3a	knowledge	apply knowledge of mathematics, science, and engineering
3b	experiment	design and conduct experiments, as well as to analyze and interpret data
3c	design	design a system, component, or process to meet economic, environmental, social or sustainable needs
3d	team	function on multidisciplinary teams
3e	solve	identify, formulate, and solve engineering problems
3f	ethical	understand your professional and ethical responsibility as an engineer
3g	communicate	communicate effectively
3h	societal	understand the impact of engineering solutions in a global, economic, environmental, and societal context
3i	learning	recognize the need for, and an ability to engage in life-long learning
3j	current events	understand how current events in society might affect engineering
3k	practice	use the techniques, skills, and modern engineering tools necessary for engineering practice

II. Measuring Engineering Self-Efficacy

The intent of the ABET Student Outcomes criterion is to define the specific behaviors that engineering students will exhibit following successful completion of the accredited education program. This works well with educational experts assessing student competency but does little to help students express where they feel they are on the road to competency. Converting these behaviors to expressions of self-efficacy offers insight to both students and educational experts as well as providing a measure of progress against the ultimate goal of competency.

A. Exisiting Measures of Engineering Self-Efficacy

The measurement of engineering self-efficacy among engineering students has generally fallen into two broad areas: academic self-efficacy and engineering task self-efficacy, as shown in TABLE II. Academic self-efficacy is defined as “a student’s belief in their capability to learn” [10] and has been shown to be an important predictor of student achievement, as higher levels of academic self-efficacy have a positive relationship with grade performance and graduation [11].

Lent, Brown and Larkin [12] were amongst the first researchers to measure academic self-efficacy within an engineering student population, measuring “strength of self-efficacy for academic milestones (AM-S)” with an 11-item scale. The AM-S scale, with items such as "rate your confidence in your ability complete the mathematics requirements for most engineering majors," showed a strong and significant relationship to both technical grade performance and persistence in engineering though graduation. Santiago and Einarson [13] adopted a broader view of academic self-efficacy among a sample of engineering and science graduate students, exploring not only academic competency (“your ability to handle the coursework”) but also academic/life interaction (“your ability to maintain a balance between your school and personal life” and “handling stress related to graduate work”). In this study, higher levels of academic self-efficacy were predicted by both student perceptions of academic preparedness and expectations about faculty/student interaction.

Dunlop [14] tackled the difficult problem of measuring academic self-efficacy in an engineering capstone class where academic requirements can vary significantly depending on the student project. This study used the Generalized Self-Efficacy scale [15] to approximate academic self-efficacy (“I can always manage to solve difficult problems if I try hard enough” and “If I am in trouble, I can usually think of a solution”) and found that over the 16-week course there was a significant increase in reported student self-efficacy. Vogt, Hocevar, and Hagedorn [16] explored the social cognitive correlates that explain differences between male and female success rates among undergraduate engineering students and identified academic self-efficacy as significant contributing variable. This was measured using an adapted version of the 7-item Self-Efficacy for Learning subscale (“I can master the skills in my major” and “I will do well in my major”) of the 44-item Motivated Strategies for Learning Questionnaire (MSLQ) [17].

As important as academic self-efficacy is to the success of an engineering student there has been a persistent interest in defining engineering success by mastery of the unique skills or tasks that define the discipline. Perhaps the first career-based assessment of engineering behavior was that of John Holland

TABLE II. ENGINEERING ACADEMIC SELF-EFFICACY AND TASK SELF-EFFICACY MEASUREMENT SCALES: EXEMPLARS

Academic Self-Efficacy	Scale Construct	# Items
Lent et al. 1986	(Academic Milestones AM-S) “ability to perform specific accomplishments critical to academic success in science and engineering majors”	11
Santiago and Einarson 1998	(Academic Self-Efficacy) “level of confidence relative to the completion of degree-related work”	10
Dunlop 2005	(Generalized Self-Efficacy) “belief that one can perform a novel or difficult tasks, or cope with adversity”	10
Vogt et al. 2007	(Motivated Strategies for Learning Questionnaire MSLQ) “self-efficacy for learning subscale”	7

Task Self-Efficacy	Scale Construct	# Items
Holland 1959, 1966	(Holland Code RI: Realistic and Investigative) "independent, stable, persistent, practical, no-nonsense, down-to-earth”	20
Purzer 2011	(Engineering Self-Efficacy) Created for a Introduction to Engineering Design course to meet course objectives. Survey items confirmed by two instructors.	16
Mamaril 2016	(Engineering Skills) 3 subscales: Experimental, Tinkering and Design	12
Gilmartin 2017	(Engineering Task Self-Efficacy) factor reduced from a longer list of items relating to engineering career outcomes	5
Chan 2017	(Competency in Generic Skills) 8 subscales: Academic and Problem-Solving, Interpersonal, Community and Citizenship, Leadership, Professional Effectiveness, Information and Communication Literacy, Critical Thinking and Self-Management	35

[18], [19] who posited that occupational satisfaction was a function of the match between an individual’s “personality” and “vocation characteristics” [20]. In Holland’s model, engineering work was best suited to the Realistic and Investigative individual (Holland code: RI) who is described as "independent, stable, persistent, practical, no-nonsense, down-to-earth” [21]. The Holland subscale for RI includes 20-items as part of a 60-items overall scale and includes behaviors like “repair household appliances,” “assemble electronic parts” and “set up and operate machines to make products” [22].

Engineering education researchers have been working to develop an all-encompassing measure of engineering task self-efficacy, often as an independent variable predicting engineering student performance. Purzer [23] developed a 16-item engineering task self-efficacy scale to study the relationship between within-group verbal exchanges, self-efficacy, and individual student achievement. This scale included items such as “confidence in your ability to explain the steps of the engineering process” and “… use Newton’s laws to solve an engineering design problem involving a falling object” and had a positive and significant relationship with student achievement in the course. Gilmartin et al. [2] developed a very brief, 5-item measure of engineering task self-efficacy as part of a national survey “to operationalize and examine innovation in the context of the broader engineering student experience.” This scale included items such as “conduct experiments, build prototypes, or construct mathematical models to develop or evaluate a design” and “analyze the operation or functional performance of a complete system” were identified from a factor analysis of a longer list of engineering task items used in Fouad and Singh’s [24] work on engineering career outcomes.

More recently, there have been efforts to develop an engineering task self-efficacy scale by starting with a broad array of engineering tasks and through factor analysis reducing the number of items to a manageable scale. Mamaril et al. [25] developed a four-factor engineering self-efficacy scale with the first factor defining engineering academic self-efficacy and the additional three factors based on the engineering task constructs of experimental skills, tinkering, and design skills. The engineering task factors started with 19-items gathered from a variety of sources and was reduced to three factors and 12-items through exploratory factor analysis. The resulting scale was considered both reliable and valid and significantly predicted GPA in academic major. Chan, Zhao, and Luk [26] developed an eight factor Generic Skills scale tested among first-year Chinese engineering students. Researchers identified 12 domains and 38 skills which they identified “as important to the engineering context” and through exploratory and confirmatory factor analysis reduced this to eight factor subscales and a total of 35-items. The scales were claimed to be both reliable and valid based on inter-item correlation and Cronbach alpha scores.

B. *Creating a Measure of Self-Efficacy*

Albert Bandura’s concept of perceived self-efficacy is an important predictor of both interests and career choice. Bandura defines perceived self-efficacy as the extent or strength of one's belief in one's own ability to complete tasks and reach goals [27], [28]. It has been shown that perceived self-efficacy is an important predictor of intended behavior in areas such as engineering academic performance [12], engineering persistence [29] and career choice [30]. In Bandura’s model, efficacy expectations moderate a person’s behavior, and in turn, this behavior is moderated by outcome expectations, leading to action toward an outcome. Bandura [31] characterizes perceived self-efficacy as agentic - the intention to make things happen by one’s action – and this makes it particularly relevant as a predictor or eventual behavior.

In the construction of self-efficacy scales Badura recommends that “items should be phrased in terms of can do rather than will do. Can is a judgment of capability; will is a statement of intention. Perceived self-efficacy is a major determinant of intention, but the two constructs are conceptually and empirically separable” [32, p. 308]. For the purposes of an engineering self-efficacy scale “can do” is interpreted a “confidence in your ability to do.” Bandura also recommends that self-efficacy scales “be tailored to activity domains and assess the multifaceted ways in which efficacy beliefs operate within the selected activity domain” [32, p. 310], which is consistent with the breadth of the ABET 3a-k criteria. This multifaceted measurement of self-efficacy is a “correlative process," per Bandura and valid and reliable self-efficacy scales should have a strong (>.80) Cronbach alpha score.

C. ABET Criteria as an Engineering Self-Efficacy Measure

While these academic and task self-efficacy scales have many advantages they also present some challenges, particularly when viewed as a possible curriculum diagnostic tool. It is clear that while the many versions of academic self-efficacy scales may predict classroom performance they do so at a very high level, often without the granularity to help understand specific areas where a student may be struggling within a specific classroom. The engineering task self-efficacy instruments, similarly, may suffer from "misplaced granularity" where each reflects what may occur only within a specific classroom experience or reflect the researchers personal vision of what defines "engineeringness."

It seems that an engineering self-efficacy scale based on the actual ABET 3.1-7 (or 3a-k) criteria would be an effective alternative to the many engineering task self-efficacy scales that exist today. The ABET criteria define the totality student outcomes for "engineeringness" established by a broad consensus of academic and industry stakeholders. The seven (or eleven) criteria offer content granularity that is at least equal if not better than what exists in current engineering task self-efficacy scales. Most importantly, a scale based on the ABET criteria is ready-made for accreditation purposes and aligns an important diagnostic tool with the ABET standards by which a program will be evaluated.

III. Research Methods

A. Participants

Data collection was conducted at six university/college locations during spring and fall semesters of 2017, resulting in 561 valid responses, as shown in TABLE III. Age of participants varied from a low of 18.1 years to a high of 30.3 years reflecting various stages of academic careers, with an average age of 19.5 years, roughly equating to a mid-second-year student. The sample is mostly comprised of mechanical engineering majors enrolled in a solid mechanics-type class. The gender split of this sample was 29% female, and 21% underrepresented minorities (URM) which is consistent with U.S. national percentages for post-secondary engineering students.

TABLE III. Sample Demographics

Location	n	Age (yrs)	% Female	% URM
A	14	20.7	21%	14%
B	49	18.1	49%	6%
C	112	19.6	51%	31%
D	123	20.1	19%	53%
E	263	19.4	22%	5%
Total	**561**	**19.5**	**29%**	**21%**

B. Data Collection and Cleaning

The survey and all measures used in the study were reviewed and authorized by the lead research university's institutional review board with concurrence from each the location. With the instructors' permission at each site, an on-line survey was administered to students participating in specific classes. There was no compensation for participation beyond a "thank you" candy bar for each member of the class regardless of survey participation. Survey questions were presented in an identical order across all survey administrations and the self-efficacy scale items were randomly presented by participant to prevent order bias.

Survey responses were accepted only from self-declared engineering majors. The data across all studies was 97.3% complete. It was determined that missing data were missing completely at random (MCAR) and multivariate imputation by chained equations (MICE) was used with predictive mean matching (PMM) and 5-iterations to complete the data set for analysis [33], [34].

C. ABET Self-Efficacy Scale Construction

The ABET-based self-efficacy scale used in this research was created using Bandura's guidelines [32] for the development of self-efficacy scales. The prompt for this scale was worded to frame the response as confidence in the respondent's ability to accomplish the behavior – *"How confident are you in your ability to ..."* The scale items were drawn from the specific ABET Criterion, either 3a-k and 3.1-7 statements. For example, ABET 3a is stated as *"an ability to apply knowledge of mathematics, science, and engineering"* and this became a survey item worded simply as *"apply knowledge of mathematics, science, and engineering."* Each respondent saw both sets of statements and responses were recorded with a slider bar that ranged from 0 ("Cannot Do At All) to 100 ("Highly Certain Can Do") with 0 as the initial setting. The individual self-efficacy items were presented in a random order that varied with each participant.

D. Analysis Procedures

Descriptive analyses were conducted by examining item means, standard deviations, and frequency distribution. The data were normally distributed with skewness and kurtosis within ± 2.0 [35] for all self-efficacy items.

Test-retest was conducted (at Location E) to confirm temporal stability, administering the same scale items to respondents twice separated by a two-week time gap and comparing results on a respondent level. At each administration respondents were assessed on their affect or "mood" which serves as a covariant measure to help explain any changes in pre-post results. Affect has long been associated with an impact on self-efficacy [36]–[39], with positive affect tending to increase self-efficacy and negative affect decreasing self-efficacy. Affect was measured "in the moment" *(how are you feeling right now?)* using the Positive and Negative Affect Schedule (PANAS) [40], specifically using a short-form, 10-item version [41]: the 10-item PANAS is comprised of 5 positive affect (PA) items and 5 negative affect (NA) items (randomly presented) rated on a scale of (1) very slightly or not at all to (5) very much. Hierarchical regression analysis comparing results from Time 1 (T1) and Time 2 (T2) controlling for changes in affect was used to confirm temporal consistency.

For the purposes of significance testing, Cohen's d [42], a measure of effect size, is used for comparison of mean

TABLE IV. ITEM AND SCALE MEANS

Code	Label	mean	SD	min	max	skew	Female	Male	*d*	URM	≠URM	*d*
Current Criteria												
3.1	solve complex	62.1	23.2	0	100	-0.39	56.8	64.3	**.33**	60.6	62.5	.08
3.2	engineering design	60.9	23.2	0	100	-0.39	57.7	62.2	.19	58.7	61.4	.12
3.3	communicate	73.6	22.0	3	100	-0.86	75.5	72.8	.12	72.0	74.0	.09
3.4	ethical	68.1	22.3	2	100	-0.64	68.2	68.1	.00	66.1	68.6	.12
3.5	team	81.4	17.4	15	100	-1.14	80.9	81.6	.04	82.4	81.1	.08
3.6	experiment	62.3	23.2	0	100	-0.48	59.1	63.6	.19	61.4	62.5	.05
3.7	learning	73.8	19.5	0	100	-0.91	73.9	73.8	.00	72.6	74.2	.08
	Total 3.1-7	**68.9**	**16.3**	**16**	**100**	**-0.56**	67.4	69.5	.13	67.7	69.2	.09
Previous Criteria												
3a	knowledge	71.2	19.1	10	100	-0.68	67.4	72.7	**.28**	70.0	71.5	.08
3b	experiment	65.9	21.7	5	100	-0.56	64.7	66.5	.08	63.4	66.6	.15
3c	design	52.7	25.3	0	100	-0.12	48.6	54.3	**.23**	51.0	53.1	.08
3d	team	76.1	19.4	10	100	-0.96	76.2	76.1	.01	75.6	76.3	.04
3e	solve	65.2	22.1	0	100	-0.58	59.2	67.7	.39	61.4	66.2	**.22**
3f	ethical	73.9	19.5	5	100	-0.71	73.3	74.1	.04	71.4	74.5	.16
3g	communicate	79.6	19.0	0	100	-1.18	79.7	79.5	.01	77.8	80.1	.12
3h	societal	71.2	20.6	0	100	-0.71	70.7	71.5	.04	68.9	71.9	.14
3i	learning	80.7	18.8	0	100	-1.27	79.5	81.2	.09	78.5	81.3	.15
3j	current events	65.5	23.0	0	100	-0.53	65.6	65.5	.00	62.9	66.2	.15
3k	practice	63.5	22.8	0	100	-0.50	57.6	65.9	**.37**	62.1	63.9	.08
	Total 3a-k	**69.6**	**15.3**	**14**	**100**	**-0.58**	67.5	70.5	.19	67.5	70.1	.17

n = 561, *d* = Cohen's *d* (<.20 no difference, .20 - .49 small difference, .50 - .79 medium difference, .80+ large difference)

difference with d > .20 considered a small significant difference, d > .50 a medium significant difference and d > .80 a large significant difference. Significance testing for correlations and linear regression uses probability-value (p-value) where p < .05 is considered significant. Data analysis was completed in R using a variety of analysis packages [43]–[45].

IV. Results

A. Item and Scale Means

Means for items on the ABET 3a-k and 3.1-7 scales are shown in TABLE IV. Items responses were across the full range (0-100) and normally distributed (skew ± 2.0). For the 3a-k scale, item means ranged from a low of 52.7 *("design a system …")* to a high of 80.7 (*"… engage in life-long learning")*. For the 3.1-7 scale, item means ranged from a low of 60.9 (*"apply engineering design …"*) to a high of 81.4 (*"function effectively on a team …*).

There we just a few, small significant differences in item means between males and females, and URM and non-URM students. For example, males had higher scores than females on item 3a *"knowledge"* (*d* = .28), item 3c "design" (Cohen's *d* = .23), item 3k *"practice"* (*d* = .37) and item 3.1 *"solve complex"* (*d* = .33) although this did not translate into a significant difference on the entire scale. Among URM students, there was a lower score on item 3e *"solve"* (*d* =.22) and this did not translate into a significant difference on the entire scale.

B. Scale Mean Comparison

The underlying construct measured by the current seven-item ABET 3.1-7 scale is virtually identical to the previous eleven-item ABET 3a-k scale. The total scale means were 69.6 for the 3a-k scale and 68.9 for the 3.1-7 scale. These means were not statistically different (*d* = .13). The Pearson r correlation coefficient between the means of the two scales is a robust .94 and regression between the two scales showed that one scale explained over 89% of the variance of the other scale (adjusted r^2 = .89, F(1,559) = 4582, p < .00).

C. Scale Reliability

The psychometric reliability of both the 3a-k and 3.1-7 scales was measured using three techniques. First, using Cronbach alpha eleven-item 3a-k scale had a strong .91 alpha while the seven-item 3.1-7 scale had a good .87 alpha indicating acceptable internal reliability for each scale. Second, a split half reliability measure was calculated and resulted in a Guttman λ4 score of .89 for the 3a-k scale and .94 for the 3.1-7 scale, where > .80 is considered acceptable [46], [47]. Finally, a test-retest was conducted on both scales (at Location E, n = 211), as shown in TABLE V. Both scales had acceptable test-retest results, particularly the 3.1-7 scale which had a correlation of .86 between T1 and T2 scores, and after controlling for net changes in affect, had an adjusted r^2 of .76 (F(2,208) = 303, p < .00), which is above the > .70 threshold considered acceptable for test-retest performance.

TABLE V. TEST-RETEST RESULTS

	Scale Mean		Pearson	Regression
Scale	**T1**	**T2**	***r***	**adjusted r^2**
3.1-7	66.8	67.1	.86	.76
3a-k	67.4	68.9	.82	.69

V. Additional Observations

A. Hierarchical Clustering

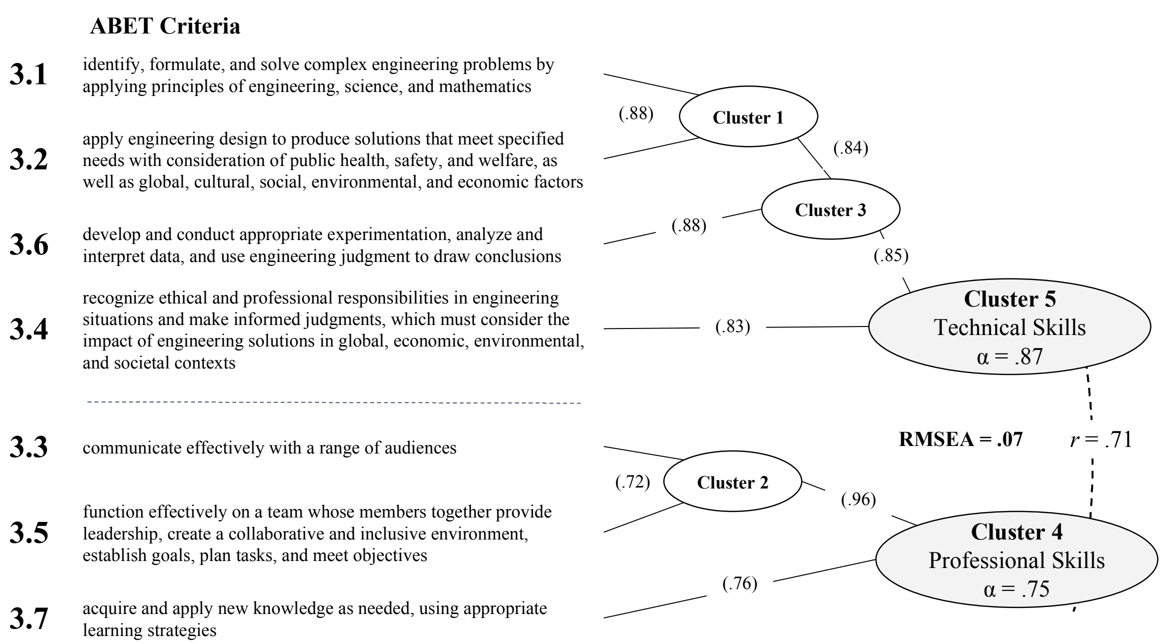
The ABET Criterion 3 can be used to help guide the development and delivery of educational resources to engineering students. To aid in this effort, it would be helpful to know it there was some way to organize the criteria that aligned with the way students this about these outcomes. The ABET criteria in this study (n = 561) was subjected to an exploratory factor analysis (EFA) to determine if there is an underlying structure to the way students think of these criteria. The EFA, using Promax rotation, showed that the criteria grouped into two factors, with items 3.1, 3.2, 3.4 and 3.6 grouping in one factor which explained 37% of the variance in the data set and items 3.3, 3.5 and 3.7 grouping in a second factor which explains an additional 23% of the variance for a total of 60% of the variance explained.

Hierarchical clustering was used to examine the structure between each of these items, as shown in Fig 1. Items 3.1 and 3.2 where most closely aligned and formed Cluster 1, while items 3.3 and 3.5 where next most closely aligned forming Cluster 2 and so on until two clusters (Cluster 4 and Cluster 5) included all the items. This data structure has a Root Mean Square Error of Approximation (RMSEA) of .07 which is considered an acceptable fit [48], [49]. Reviewing the items in each cluster, it appears that Cluster 5 could be renamed "Technical Skills" (α = .87) while Cluster 4 appears to describe "Professional Skills" (α = .75). This high-level grouping of the ABET Criterion 3 items into Technical Skills and Professional Skills may help in the communication and implementation of this criteria into curriculum.

B. ABET 3.1-7 Short Form

An objective of this study was to explore the use of ABET criteria as a possible measure of engineering self-efficacy. It appears that the ABET 3.1-7 criteria can serve as an effective measure of overall engineering self-efficacy. However, as a survey question the ABET 3.1-7 criteria is long (7 items, 127 words, 18.1 words per item) perhaps inducing "survey fatigue" particularly when included as only part of a longer survey. It may be possible to create a "short form" version of ABET Criteria by combining the validity of 3.1-7 with the brevity of 3a-k (11 items, 107 words, 9.7 words per item) to address the uncomfortable choice of using a brief scale to measure engineering self-efficacy or using no scale at all.

As discussed earlier, the 3.1-7 and 3a-k criteria seem to measure the same underlying construct and are statistically similar. A subset of the briefer 3a-k criterion statements can be used to predict the totality of the 3.1-7 criteria through multiple regression, incrementally adding one 3a-k criteria at a time, as shown in Fig 2. For example, regressing criteria 3g (the single best fit variable) against the mean of all criteria in 3.1-7 results in a .34 model fit (adjusted r^2). Then adding criteria 3k (i.e., 3g + 3k) results in a combined .77 model fit. By progressively adding the 3a-k variables to the regression shows where the addition of variables begins to contribute only a modest improvement in overall model fit. The inflection point appears to be four criteria (3g, 3k, 3b and 3c) which combined have a .85 adjusted r^2 model fit with the mean of all 3.1-7 criteria.



(Factor Loadings), α = Cronbach's alpha, *r* = Pearson's Correlation

Fig. 1. Hierachical Cluster Analysis of the Current ABET Criterion 3 – Student Outcomes

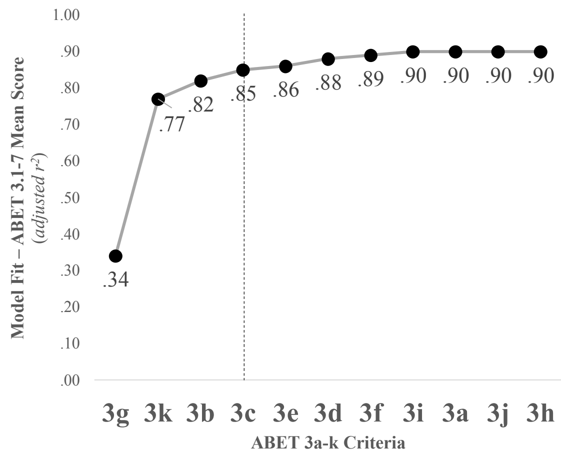


Fig. 2. Model Fit (adjusted r^2) of ABET 3.1-7 Mean by Progressively Adding ABET 3a-k Criterion in Order of Impact

The resulting short form engineering self-efficacy measure (ESE-SF) has just 4 items, 42 words with 10.5 words per item as shown in TABLE VI. The ESE-SF scale has a significant .92 Pearson correlation with the 3.1-7 scale mean and the model fit, as stated earlier, has a .85 adjusted r^2 ($F(1,559) = 3018$, $p < .00$). The ESE-SF has a .81 Cronbach alpha and a split half reliability (Guttman λ4) of .85, both indicating good internal reliability. As a result, the ESE-SF scale can be a reliable predictive indicator of the engineering self-efficacy described by 3.1-7 scale as a unitary construct.

TABLE VI. ENGINEERING SELF-EFFICACY SHORT FORM (ESE-SF)

How confident are you in your ability to ...
- communicate effectively
- use the techniques, skills, and modern engineering tools necessary for engineering practice.
- design and conduct experiments, as well as to analyze and interpret data
- design a system, component, or process to meet economic, environmental, social or sustainable needs

VI. DISCUSSION

We are encouraged by these results showing that the current ABET criteria can be transformed into a functional engineering self-efficacy scale (ESE-ABET). An early reviewer asked the obvious question … *why does the world need yet another engineering self-efficacy scale?* An engineering self-efficacy scale based on ABET criteria has two advantages. First, it has the strong face validity in that it measures the totality of engineering behavior (not a subset of behaviors or specific skill set), constructed with criteria that is based on broad-based input from both the academy and industry. Second, an engineering self-efficacy scale specifically based on ABET criteria can be a useful tool in helping organizations work toward ABET accreditation.

However, this ABET-based self-efficacy scale is not perfect. The ABET criteria were written to provide organizational guidance, not individual self-reflection, and as a result some of the items tend to be wordy and not tailored to a single person. Also, from a psychometric measurement perspective, some of the resulting 3.1-7 statements are "double barreled" in that they seem to describe two behaviors within one criteria. For example, Criteria 3.4 *"ethical solutions"* (see TABLE I) seems to be a combination of the previous 3f *"ethical"* and 3h *"societal."* The shorter statements of the previous Criteria 3a-k had more of a singular focus and made behavior measurement more precise. More research on the impact of the 3.1-7 criteria language used as a psychometric measure would help sort this out.

With these two reliable measures of "engineeringness," (the complete ESE-ABET and the ESE-SF) the future is bright with possibilities. These measures could be used as an "annual check-up" for an engineering school to help track grade level cohorts on their way toward becoming a successful workplace engineers. These measures could be used on the class-level, perhaps pre-and-post, to determine if certain activities had a positive impact on engineering self-efficacy. As another possibility, a measure of engineering self-efficacy might be able to help differentiate how certain groups of students (e.g., ethnic diverse, first generation or women) prosper in some environments and fall behind in others. As a tool, ESE-ABET and ESE-SF open the door to entirely new areas of learning in the future.

ACKNOWLEDGMENT

We would like to thank the ABET Engineering Accreditation Commission (EAC) for allowing the use of their newly developed Criterion 3 in this research before it was broadly available and for their continuing support of this research effort. We would also like to thank our research colleagues who made their classes available for study and provided helpful advice in the process - Dr. Allison Godwin of Purdue University, Dr. Micah Lande of Arizona State University, Dr. Bob Rice of UC-Merced and Dr. Bob Witt of the University of Wisconsin - Madison. Finally, a heart-felt thank you to the engineering students who have given their time and attention to this project.

REFERENCES

[1] U. S. National Academy of Engineering, *The engineer of 2020: visions of engineering in the new century*. National Academies Press Washington, DC, 2004.
[2] S. K. Gilmartin *et al.*, "Designing a Longitudinal Study of Engineering Students' Innovation and Engineering Interests and Plans: The Engineering Majors Survey Project. EMS 1.0 and 2.0 Technical Report.," Stanford University Designing Education Lab, Stanford, CA, Technical Report, Feb. 2017.
[3] "ABET | Constitution," Jun-2017. [Online]. Available: http://www.abet.org/about-abet/governance/board-of-directors/constitution/.
[4] "ABET | History," Jun-2017. [Online]. Available: http://www.abet.org/about-abet/history/.
[5] "ABET | Accreditation Criteria," Jun-2017. .
[6] Engineering Accreditation Commission, *Engineering Criteria 2000: Criteria for Accrediting Programs in Engineering in the United States*, 3rd edition. Baltimore,

MD: Accreditation Board for Engineering and Technology (ABET), 1997.
[7] J. F. Volkwein, L. R. Lattuca, P. Terenzini, L. Strauss, and J. Sukhbaatar, "Engineering Change: A Study of the Impact of EC2000," *International Journal of Engineering Education*, vol. 20, no. 3, pp. 318–328, 2004.
[8] J. F. Volkwein, L. R. Lattuca, B. J. Harper, and R. J. Domingo, "Measuring the Impact of Professional Accreditation on Student Experiences and Learning Outcomes," *Research in Higher Education*, vol. 48, no. 2, pp. 251–282, 2007.
[9] "ABET | Criterion 3 Revision Timeline," Jun-2017. [Online]. Available: http://www.abet.org/criterion-3-revision-timeline/.
[10] B. J. Zimmerman, A. Bandura, and M. Martinez-Pons, "Self-Motivation for Academic Attainment: The Role of Self-Efficacy Beliefs and Personal Goal Setting," *American Educational Research Journal*, vol. 29, no. 3, pp. 663–676, 1992.
[11] F. Pajares, "Self-Efficacy Beliefs in Academic Settings," *Review of Educational Research*, vol. 66, no. 4, pp. 543–578, 1996.
[12] R. W. Lent, S. D. Brown, and K. C. Larkin, "Self-Efficacy in the Prediction of Academic Performance and Perceived Career Options," *Journal of Counseling Psychology*, vol. 33, no. 3, pp. 265–269, 1986.
[13] A. M. Santiago and M. K. Einarson, "Background Characteristics as Predictors of Academic Self-Confidence and Academic Self-Efficacy Among Graduate Science and Engineering Students," *Research in Higher Education*, vol. 39, no. 2, pp. 163–198, Apr. 1998.
[14] J. C. Dunlap, "Problem-Based Learning and Self-Efficacy: How a Capstone Course Prepares Students for a Profession," *Educational Technology Research and Development*, vol. 53, no. 1, pp. 65–83, 2005.
[15] R. Schwarzer and M. Jerusalem, "Generalized Self-Efficacy scale," in *Measures in Health Psychology: A Users Portfolio. Causal and Control Beliefs*, J. Weinman, S. Wright, and M. J. Azur, Eds. Windsor, UK: NFER-NELSON, 1995, pp. 35–37.
[16] C. M. Vogt, D. Hocevar, and L. S. Hagedorn, "A Social Cognitive Construct Validation: Determining Women's and Men's Success in Engineering Programs," *The Journal of Higher Education*, vol. 78, no. 3, pp. 337–364, 2007.
[17] P. R. Pintrich, D. A. F. Smith, T. Garcia, and W. J. McKeach, "A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)," Office of Educational Research and Improvement, Ann Arbor, MI, NCRIPTAL-91-B-004, 1991.
[18] J. L. Holland, "A theory of vocational choice.," *Journal of counseling psychology*, vol. 6, no. 1, p. 35, 1959.
[19] J. L. Holland, *The psychology of vocational choice: A theory of personality types and model environments*. Oxford, England: Blaisdell Publishing COmpany, 1966.
[20] M. M. Nauta, J. H. Kahn, J. W. Angell, and E. A. Cantarelli, "Identifying the antecedent in the relation between career interests and self-efficacy: Is it one, the other, or both?," *Journal of Counseling Psychology*, vol. 49, no. 3, p. 290, 2002.
[21] G. D. Gottfredson and J. L. Holland, *Dictionary of Holland occupational codes*. Psychological Assessment Resources Inc, 1996.
[22] US Department of Labor - Employment and Training Administration, "O*Net Resource Center - Interest Profiler," Jun-2017. [Online]. Available: https://www.onetcenter.org/IP.html.
[23] S. Purzer, "The Relationship Between Team Discourse, Self-Efficacy, and Individual Achievement: A Sequential Mixed-Methods Study," *Journal of Engineering Education*, vol. 100, no. 4, pp. 655–679, 2011.
[24] N. A. Fouad, R. Singh, M. E. Fitzpatrick, and J. P. Liu, "Stemming the Tide: Why Women Leave Engineering," University of Wisconsin, Milwaukee, WI, 2011.
[25] N. A. Mamaril, E. L. Usher, C. R. Li, D. R. Economy, and M. S. Kennedy, "Measuring Undergraduate Students' Engineering Self-Efficacy: A Validation Study," *J. Eng. Educ.*, vol. 105, no. 2, pp. 366–395, Apr. 2016.
[26] C. K. Chan, Y. Zhao, and L. Y. Luk, "A Validated and Reliable Instrument Investigating Engineering Students' Perceptions of Competency in Generic Skills," *Journal of Engineering Education*, vol. 106, no. 2, pp. 299–325, 2017.
[27] A. Bandura, "Self-Efficacy: Toward a Unifying Theory of Behavioral Change.," *Psychological review*, vol. 84, no. 2, p. 191, 1977.
[28] A. Bandura, *Self-Efficacy: The Exercise of Control*, 1st edition. New York: Worth Publishers, 1997.
[29] J. P. Concannon and L. H. Barrow, "Men's and Women's Intentions to Persist in Undergraduate Engineering Degree Programs," *Journal of Science Education and Technology*, vol. 19, no. 2, pp. 133–145, 2010.
[30] R. W. Lent, H.-B. Sheu, D. Singley, J. A. Schmidt, L. C. Schmidt, and C. S. Gloster, "Longitudinal Relations of Self-Efficacy to Outcome Expectations, Interests, and Major Choice Goals in Engineering Students," *Journal of Vocational Behavior*, vol. 73, no. 2, pp. 328–335, 2008.
[31] A. Bandura, "Social Cognitive Theory: An Agentic Perspective," *Annual review of psychology*, vol. 52, no. 1, pp. 1–26, 2001.
[32] A. Bandura, "Guide for Constructing Self-Efficacy Scales," in *Self-efficacy beliefs of adolescents*, vol. 5, F. Pajares and T. C. Urdan, Eds. Information Age Publishing, 2006, pp. 307–337.
[33] C. A. Manly and R. S. Wells, "Reporting the Use of Multiple Imputation for Missing Data in Higher Education Research," *Research in Higher Education*, vol. 56, no. 4, pp. 397–409, 2015.
[34] S. van Buuren and K. Groothuis-Oudshoorn, "mice: Multivariate Imputation by Chained Equations in R," *Journal of Statistical Software*, vol. 45, no. 3, pp. 1–67, Dec. 2011.

[35] D. R. Krathwohl, *Methods of Educational and Social Science Research: The Logic of Methods*, 3rd ed. Waveland Press, 2009.
[36] D. J. Kavanagh and G. H. Bower, "Mood and self-efficacy: Impact of joy and sadness on perceived capabilities," *Cognitive Therapy and Research*, vol. 9, no. 5, pp. 507–525, 1985.
[37] R. A. Baron, "Environmentally Induced Positive Affect: Its Impact on Self-Efficacy, Task Performance, Negotiation, and Conflict," *Journal of Applied Social Psychology*, vol. 20, no. 5, pp. 368–384, 1990.
[38] J. B. Thatcher and P. L. Perrewe, "An empirical examination of individual traits as antecedents to computer anxiety and computer self-efficacy," *Mis Quarterly*, pp. 381–396, 2002.
[39] J. B. Vancouver, C. M. Thompson, E. C. Tischner, and D. J. Putka, "Two studies examining the negative effect of self-efficacy on performance.," *Journal of Applied Psychology*, vol. 87, no. 3, p. 506, 2002.
[40] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: The PANAS scales," *Journal of Personality and Social Psychology*, vol. 54, no. 6, pp. 1063–1070, 1988.
[41] E. R. Thompson, "Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS)," *Journal of Cross-Cultural Psychology*, vol. 38, no. 2, pp. 227–242, Mar. 2007.
[42] J. Cohen, "A power primer.," *Psychological bulletin*, vol. 112, no. 1, p. 155, 1992.
[43] R Core Team, *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2015.
[44] W. Revelle, "Hierarchical Cluster Analysis and the Internal Structure of Tests," *Multivariate Behavioral Research*, vol. 14, no. 1, pp. 57–74, 1979.
[45] Y. Rosseel, *lavaan: An R Package for Structural Equation Modeling*. 2012.
[46] L. Guttman, "A Basis for Analyzing Test-Retest Reliability," *Psychometrika*, vol. 10, no. 4, pp. 255–282, 1945.
[47] J. C. Callender and H. G. Osburn, "An Empirical Comparison of Coefficient Alpha, Guttman's Lambda-2, and Msplit Maximized Split-Half Reliability Estimates," *Journal of Educational Measurement*, vol. 16, no. 2, pp. 89–99, 1979.
[48] J. B. Schreiber, A. Nora, F. K. Stage, E. A. Barlow, and J. King, "Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review," *The Journal of Educational Research*, vol. 99, no. 6, pp. 323–338, Jul. 2006.
[49] L. Hu and P. M. Bentler, "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural equation modeling: a multidisciplinary journal*, vol. 6, no. 1, pp. 1–55, 1999.