Complementing Assessment Processes with Standardized Tests: A Work in Progress

Amparo Camacho

***Abstract*—**ABET accredited programs must assess the development of student learning outcomes (SOs) in engineering programs. Different institutions implement different strategies for this assessment, and they are usually designed "in house." This paper presents a proposal for including standardized tests to complement the ABET assessment model in an engineering college made up of six distinct engineering programs. The engineering college formulated a model of quality assurance in education to be implemented throughout the six engineering programs to regularly assess and evaluate the achievement of SOs in each program offered. The model uses diverse techniques and sources of data to assess student performance and to implement actions of improvement based on the results of this assessment. The model is called "Assessment Process Model" and it includes SOs A through K, as defined by ABET. SOs can be divided into two categories: "hard skills" and "professional skills" (soft skills). The first includes abilities, such as: applying knowledge of mathematics, science, and engineering and designing and conducting experiments, as well as analyzing and interpreting data. The second category, "professional skills", includes communicating effectively, and understanding professional and ethnical responsibility. Within the Assessment Process Model, various tools were used to assess SOs, related to both "hard" as well as "soft" skills. The assessment tools designed included: rubrics, surveys, questionnaires, and portfolios. In addition to these instruments, the Engineering College decided to use tools that systematically gather consistent quantitative data. For this reason, an in-house exam was designed and implemented, based on the curriculum of each program. Even though this exam was administered during various academic periods, it is not currently considered standardized. In 2017, the Engineering College included three standardized tests: one to assess mathematical and scientific reasoning and two more to assess reading and writing abilities. With these exams, the college hopes to obtain complementary information that can help better measure the development of both hard and soft skills of students in the different engineering programs. In the first semester of 2017, the three exams were given to three sample groups of students from the six different engineering programs. Students in the sample groups were either from the first, fifth, and tenth semester cohorts. At the time of submission of this paper, the engineering college has descriptive statistical data and is working with various statisticians to have a more in-depth and detailed analysis of the sample group of students' achievement on the three exams. The overall objective of including standardized exams in the assessment model is to identify more precisely the least developed SOs in order to define and implement educational strategies necessary for students to achieve them in each engineering program.

***Keywords*—**Assessment, hard skills, soft skills, standardized tests.

A. Camacho is with Universidad del Norte, Barranquilla, Atl 81007 Colombia (phone: +57+5301-3643203; e-mail: acamacho@uninorte.edu.co).

I. Introduction

HIGHER education has been transforming. Nowadays, it intends to generate significant education for students as future professionals, not only in knowledge but also in the development of abilities and competencies. Within these desired abilities, the development of mathematical and scientific thinking and the acquisition of communicative skills are fundamental to a comprehensive education. Trifone says that the development of scientific reasoning takes an important place in the teaching-learning process [1]. The development of formal thinking helps students to respond effectively to problem situations, allowing significant learning. On the other hand, communication skills help students to acquire knowledge and to express a critical position on a specific subject.

In the past, there has been research about cognitive development and the distinct types of reasoning utilized by students. One study, conducted by Fleming & Morning, shows that university programs directed towards cognitive development help decrease withdrawal rates [1]. The programming focused on the cognitive development of the student. Likewise, courses related to communicative skills currently hold more weight in the academic programs for each distinct undergraduate discipline.

The paper shows preliminary results of applying standardized tests to measure development of different skills. Participating students belong to six different engineering programs (Civil Engineering, Electrical Engineering, Electronic Engineering, Industrial Engineering, Mechanical Engineering and Systems Engineering). The sample was taken from students in their first, fifth, and tenth semesters of their respective disciplines.

The objective of these measurements is to obtain results that can complement the internal assessment results of each of the six engineering programs. Therefore, the results may help to precise the abilities and knowledge that require improvement strategies. The paper employs the Student Outcomes from ABET (Accreditation Board of Engineering and Technology).

This paper presents a conceptual framework for the tests applied, particularly for the Lawson test and the assessment model developed in the engineering college, the methodology followed for the test application, and some important results. The process of integrating the assessment tools utilized with those newly resulting from the applied tests will be illustrated. With the results in mind, recommendations and future work to be developed within the existing model are presented.

II. Conceptual Framework

Scientific reasoning plays an important role in the teaching science courses. Lawson [2] identifies seven types of reasoning to understand which are the underlying processes in problem solving and how these facilitate the student's acquisition of information as well as cognitive development. The seven types of reasoning are: Conservation of mass and volume, proportional thinking, probabilistic thinking, identification and control of variables, correlational thinking, combinatorial thinking and hypothetical-deductive thinking.

Proportional thinking allows one to give a solution to the situations that can be resolved with consideration for the 'rationale' between two numbers and relate that with a third number to obtain an unknown number (…) representing a particular context for multiplicative situations [3]. Meanwhile, one can understand variable control as the capacity to identify and examine the effects of a variable. From this, one can identify the dynamic between the behavior of diverse variable sets, the effects they can have on one another and how they can remain constant [4]. Likewise, correlational thinking is understood as the model of thinking which the individual utilizes to define the relationship or reciprocity between forces or variables. This model of thinking is fundamental for identifying relationships as well as for making predictions and scientific investigations; therefore, it constitutes the base for the development of formal thinking [5]. Finally, combinatorial thinking refers to "a person's ability to conceive and systematically organize all the possibilities and dimensions that can interact as elements and causes of a problem or complex even" [6].

Once each type of thinking was identified, Lawson developed a test named Lawson's test of formal thinking, through which formal or scientific thinking can be evaluated [7]. The student's performance on the test then identifies their level of formal or scientific thinking. On the other hand, engineers need to improve their communicative skills. These skills allow the individual to develop in the different social contexts in which they interact:

> The compilation of knowledge, abilities, and skills that require the adequate, correct, coherent, and aesthetic use of oral as well as written text (comprehension and expression, analysis and synthesis, identification, comparison, creation, recreation… of messages), centering in listening and speaking, reading and writing in a competent way [8].

The engineering college uses an assessment process based in the guidelines provided by ABET, including "one or more processes that identify, collect, and prepare data to evaluate the attainment of student outcomes. Effective assessment uses relevant, direct, indirect, quantitative and qualitative measures as appropriate to the outcome being measured" [9]. During this process, rubrics were designed and utilized as an assessment tool.

The rubrics allow for assessment based on the descriptions of the performances to be achieved by a student in a learning process; they allow the identification of the performance dimensions that are being evaluated and taught and what is expected of students [10].

III. Methodology

This section shows the methodology developed to measure the achievement of the Student Outcomes from the ABET EC 2000 model and the methodology used for the application of the standardized tests. The objective will be to integrate the tests to the existing quality assurance model for engineering students' education.

A. Assessment Process Model

The assessment process is founded in the accrediting for Engineering programs criteria defined by ABET [11] in their international engineering program accreditation guidelines.

The College of Engineering created and implemented a three-nestled loops model, by means of developing the assessment process for the six engineering programs. This process constitutes the core of quality assurance for the education of future engineers. The process' focus is the continued improvement of each iteration of the cycles over time, measured in academic periods, either semesters or years.

The objective of the process is to measure and evaluate the achievement of the competencies or skills defined in the graduation profile of each of the six programs and to implement the pertinent improvement measures. The skills are based in those proposed by ABET, such as the Student Outcomes A through K [9]. It is worth noting that this model seeks the overall improvement of each engineering program relative to the quality of education given by the same; it does not intend to monitor individually the students' education. Fig. 1 illustrates the model.

- The inner loop measures the development of the learning outcomes of each of the courses that contribute to the development of some of the student outcomes of the program. Actions for improvement that are taken in the course should contribute to the development of the respective student outcomes.
- The middle loop measures the global achievement level of the student outcomes for the program. Actions for improvement taken in this cycle should contribute to the betterment of the program globally, in terms of the student outcomes.
- The outer loop allows for the revision of the program educational objectives in terms of the relevance of the program and its effect in relation to the professional field. Actions for improvement taken in this cycle should contribute to achieving the program educational objectives.

Different tools were used to obtain the required data for each measurement cycle. For the objective of this paper, the rubrics were highlighted as the tool that allowed for the measurement of the level of achievement for each one of the student outcomes of the program (middle loop).

The performance criteria of the rubrics consist of global learning objectives for the entire engineering program, as opposed to specific learning outcomes for particular courses. For this reason, the rubrics can be applied universally [12].

To minimize the variability in the measurement of the student outcomes, each academic program applied the rubrics to evaluate specific activities [13], [14].

Each program self-defined a process to periodically assess the student outcomes. The processes consisted of planning, design, evaluation and recommendations for improvement.

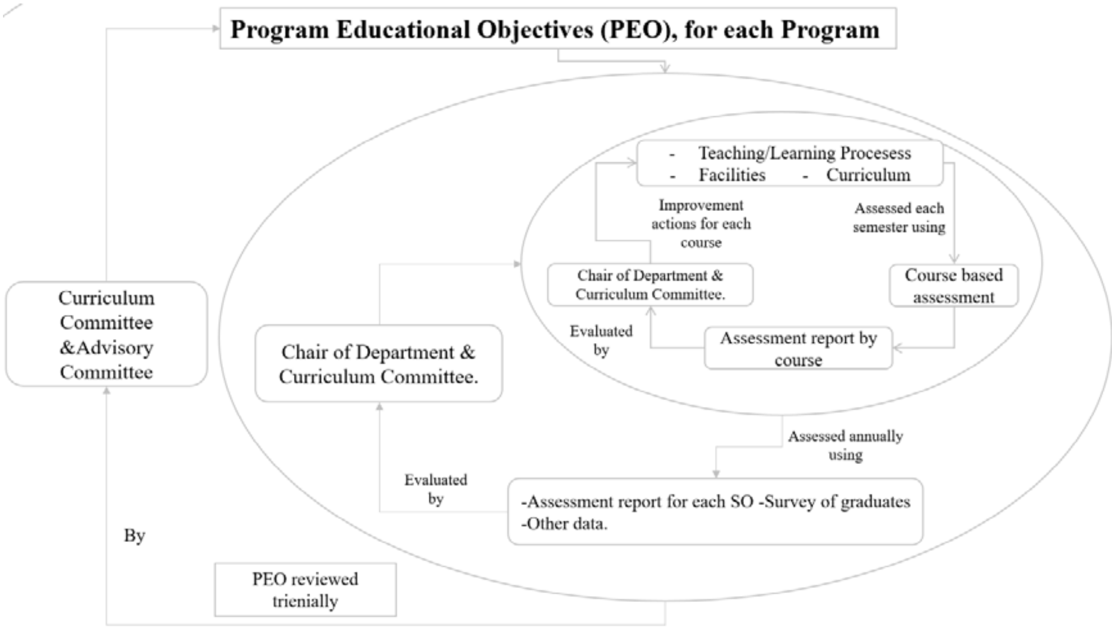

Fig. 1 Assessment Process Model

B. Standardized Tests

To obtain more precise information about the development of student outcomes for each program and to supplement that information, it was decided that starting in 2017, three standardized tests would be applied: Lawson's test, a writing test, and a reading test.

C. Description of the Sample

Data were provided from the student population of the six engineering programs from three courses, one of each from the first, fifth, and tenth semester. The student sample was random and the sample sizes are observed in Table I.

TABLE I
DESCRIPTION OF THE SAMPLES BY TEST

Semester	Lawson's test	Reading test	Writing test
1	486	155	87
5	277	43	36
10	100	41	57
Total	**863**	**239**	**180**

Lawson's test had a total sample of 863 students, male and female, belonging to the undergraduate programs of Electronic Engineering (69), Electric Engineering (91), Sytems Engineering (129), Civil Engineering (161), Industrial Engineering (227) and Mechanical Engineering (186).

D. Instruments

Three tests were utilized: Lawson's test of formal reasoning for the scientific thinking, a reading diagnostic test, and a writing diagnostic test which were developed by the university languages department based on standardized tests that measure the stated competencies. Lawson's test is described below, as the other two test results have not been completed yet.

Lawson's test for scientific reasoning fundamentally takes into account the theory of cognitive development conception proposed by Inhelder and Piaget [5], where individuals are scientists in search of knowledge. The format and style of the test questions is multiple-choice with one correct answer. This test has a distinctive aspect in which each two test items are dependent on each other in order to be marked correct. As it is necessary to answer an item which responds to a problem situation, the next item requires the test taker to choose the reasoning or argument that explains why they chose the first test item [3].

E. Data Analysis

To analyze the test results initially, a Kolmogorov–Smirnov test (KS test) was run, from the goodness of fit, which served to contrast the null hypothesis that the behavior of a variable will adjust itself to a determined theoretical distribution of probability. The results of the KS test indicate that the normality hypothesis with a critical level of $p<0.05$ was rejected, and it concludes that the variable rankings are not adjusted to a normal distribution. Thus, non-parametric statistics should be utilized to analyze the data. Due to the results above, a Kruskal–Wallis H test was run.

IV. RESULTS

In this section, the results of the standardized tests are integrated into the assessment process and an example of how to integrate them is shown. The results of the tests are integrated with the assessment process, in the model presented in Fig. 2.

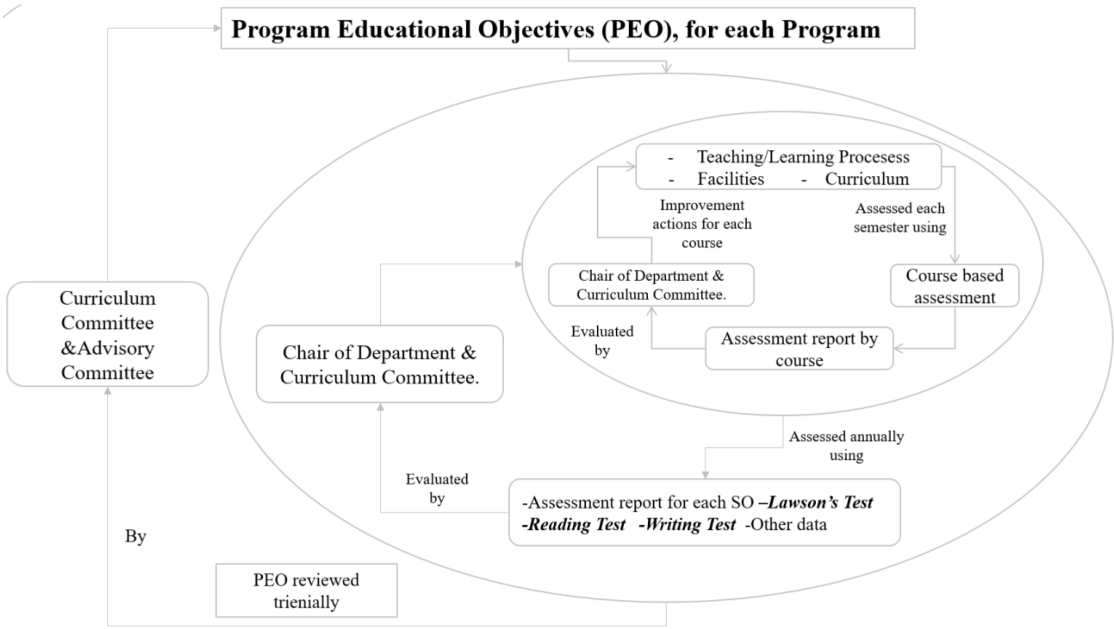

Fig. 2 Restructured Assessment Process Model

TABLE II
RUBRIC FOR S.O. A - INDUSTRIAL ENGINEERING

SO (a): "An ability to apply knowledge of mathematics, science, and engineering." [15]

Summary Table	Unsatisfactory	Developing	Satisfactory	Exemplary
1. Apply mathematical concepts. *Correction *Coherence	Most of the necessary concepts for the problem's solution are applied in an incorrect or incoherent manner.	Some of the necessary concepts for the problem's solution are applied in an incorrect or incoherent manner.	Most of the necessary concepts for the problem's solution are applied in a correct or coherent manner.	All of the necessary concepts for the problem's solution are applied in a correct or coherent manner.
2. Apply science concepts. *Correction *Coherence	Most of the necessary concepts for the problem's solution are applied in an incorrect or incoherent manner.	Some of the necessary concepts for the problem's solution are applied in an incorrect or incoherent manner.	Most of the necessary concepts for the problem's solution are applied in a correct or coherent manner.	All of the necessary concepts for the problem's solution are applied in a correct or coherent manner.
3. Apply engineering concepts. *Correction *Coherence	Most of the necessary concepts for the problem's solution are applied in an incorrect or incoherent manner.	Some of the necessary concepts for the problem's solution are applied in an incorrect or incoherent manner.	Most of the necessary concepts for the problem's solution are applied in a correct or coherent manner.	All of the necessary concepts for the problem's solution are applied in a correct or coherent manner.

TABLE III
RESULTS FOR S.O. (A)

SUMMARY TABLE	Unsatisfactory	Developing	Satisfactory	Exemplary
1. Apply mathematical concepts. *Correction *Coherence	12,50%	25,00%	25,00%	37,50%
2. Apply science concepts. *Correction *Coherence	12,50%	25,00%	50,00%	25,00%
3. Apply engineering concepts. *Correction *Coherence	12,50%	25,00%	25,00%	37,50%

At the time of writing this paper, the results of the Lawson's test have been completed with the corresponding descriptive analysis. The reading and writing test results have been partially completed.

This section will focus solely on the Lawson's test highlighting some of the results. Furthermore, an illustration of how to complement the assessment model with Lawson's test results will be shown. As an example, this section shows results of assessment of Student Outcome A (defined by ABET) for the Industrial Engineering program. Student Outcome A was selected because for its development as it requires some of the abilities measured on the Lawson's test. The rubric that measures Student Outcome A for the Industrial Engineering program is presented in Table II.

The rubric above was applied to a sample of 81 students distributed in courses at the halfway point of their undergraduate program (2 courses) and at the end of their program (2 courses). The global results obtained are shown in Table III. From the results, it can be observed that the performance demonstrated in the performance indicators of the rubric show that the students have a better performance in the application of scientific concepts for solving problems than in the application of mathematical or engineering concepts.

The results of the Lawson's test, applied to a sample of 227 students in the Industrial Engineering program, distributed by first semester (60), fifth semester (118) and tenth semester (49), are shown in Table IV.

TABLE IV
LAWSON'S TEST RESULT

	M.1	DS.	M.2	DS	M. 3	DS	Chi-squared	df	Sig. Asymptotic
Conservation of Mass and Volume	69,17	26,187	78,81	28,798	84,69	25,422	10,886	2	0,004
Proportional Thinking	30	35,891	48,31	38,637	55,1	37,144	13,548	2	0,001
Control of Variables	26,67	22,011	42,8	25,128	53,06	29,151	26,611	2	0
Probabilistic Thinking	40,83	41,672	52,54	44,264	76,53	34,007	18,854	2	0
Correlational Thinking	36,67	48,596	67,80	46,925	65,31	48,093	16,881	2	0
Combinatorial Thinking	28,33	45,442	25,42	43,729	24,49	43,448	0,247	2	0,884
Hypothetico-deductive Reasoning	37,6389	19,737	51,977	20,5190	61,224	22,149	30,338	2	0

The results obtained in the application of the Lawson's test indicate that there is a significant difference between the averages of the results obtained by students in the three semesters, in the majority of the abilities that are measured. These significant differences, marked by the increase in averages, can be interpreted as the strengthening of the abilities and the acquisition of new knowledge because of the students' educational process.

The results obtained by the Lawson's test allowed for a detailed analysis of the required abilities for the development of Student Outcome A, given the performance indicators of the rubric are global. In this way, it is possible to identify the abilities assessed in Lawson's test needed for each of the performance indicators assessed in the rubric. From there, it can be determined precisely which abilities have low performance for each performance indicator and relevant actions for improvement can be taken. Additionally, the results of the Lawson's test to measure the beginning, middle and end of a student's undergraduate program provides a more complete vision of the performance of the scientific reasoning skills which contributes to the development of assessment in this way with more complete results. Thus, both Lawson's test and the reading and writing tests can complement the results obtained by the Student Outcomes rubrics.

V. PRELIMINARY CONCLUSIONS

Given that at the time of writing the paper the required results of the three tests were not available to establish if there exist any correlations, definitive conclusions cannot be formed. The only preliminary observations are related to the application of the Lawson's test.

It was observed with the example shown that there is an evolution throughout of the undergraduate program, in terms of better performance on the test for the majority of the measured abilities. These results complement those obtained with the rubric that measured Student Outcome A, which gives greater precision to the abilities evaluated in the global rubric categories. In addition, the Lawson's test was given to a sample of students in their first year of the undergraduate program, whilst the rubric was used to evaluate courses in the middle and end of the program. This allowed for establishing a complete vision of the performance of this ability in three important semesters during the undergraduate program: the beginning, the middle, and the end. The majority of the rubrics utilized to measure the Student Outcomes of the six engineering programs took samples from students located in at least their fourth semester. Therefore, Lawson's test provides data for a more complete analysis.

With the complete results, there will be a statistical analysis, required by inferential statistics, to determine the adequate strategies for the improvement and/or strengthening of the Student Outcomes of each program.

VI. FUTURE WORK

The importance of the integration of the three tests in the assessment process resides in providing reliable and precise information to improve the quality of education for the students in the College of Engineering. For this reason, in 2018, the three tests will continue to be implemented with the three established cohorts of students, and the results will be integrated with the rubric results. These results measure the Student Outcomes that require the abilities evaluated on the three tests.

For 2018, the results of the three tests will be integrated to the following Student Outcomes:

(a) "An ability to apply knowledge of mathematics, science, and engineering." [15]
(b) "An ability to design and conduct experiments, as well as to analyze and interpret data." [15]
(c) "An ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability." [15]
(d) "An ability to identify, formulate, and solve engineering problems." [15]
(e) "An ability to communicate effectively." [15]
(f) "An ability to use the techniques, skills, and modern engineering tools necessary for engineering practice." [15]

With the preliminary results obtained in 2017, a pedagogic strategy was formulated focusing on the strengthening of the weak abilities identified by the results of the Lawson's test. This new strategy will be implemented in the first semester of 2018.

ACKNOWLEDGMENT

The author thanks the director of the Center for teacher excellence, Eulises Dominguez, CEDU by its acronym in Spanish, Gina Camargo and Roxana Quintero for the information provided. The author also thanks Romario Peña for the support in data handling.

REFERENCES

[1] J. Fleming & C. Morning, "Correlates of the SAT in Minority Engineering Students: An Exploratory Study," (online) *The Journal of Higher Education*, vol. 69, no. 1, pp. 89-108. Jan 1998. Available: http://www.jstor.org/stable/2649183

[2] A. Lawson, H. Adi & R. Karplus, "Development of Correlational Reasoning in Secondary Schools: Do Biology Courses Make a Difference?" (online) *The American Biology Teacher,* vol. 41, no. 7, pp. 420-425, Oct. 1979. Available: http://doi.org/10.2307/4446678

[3] C. Fernández & S. Linares, "De la Estructura Aditiva a la Multiplicativa: Efecto de Dos Variables en el Desarrollo del Razonamiento Proporcional", (online) *Infancia y Aporendizaje,* vol. 34, no. 1, pp. 67-80, Jan. 2011. Available: https://doi.org/10.1174/021037011794390111

[4] D. Kuhn, K. Iordanou, M. Pease & C. Wirkala, "Beyond Control of Variables: What Needs to Develop to Achieve Skilled Scientific thinking?" (online) *Cognitive Development,* vol. 23, no. 4, pp. 435-451, Dec. 2008. Available: https://doi.org/10.1016/j.cogdev.2008.09.006

[5] B. Inhelder & J. Paiget, "De la Lógica del Niño a la Lógica del Adolescente", Barcelona: España Edisiones Paidós Ibérica, 1955.

[6] J. Morales, *Operaciones Combinatorias en Adolescentes y Jóvenes Universitarios* (online). Peru: Pontificia Universidad Católica del Perú, 2013. Available: http://tesis.pucp.edu.pe/repositorio/handle/123456789/4703

[7] J. Trifone, "The Test of Logical Thinking: Applications for Teaching and Placing Science Students, *The American Biology Teacher,* vol. 49, no. 8, pp. 411-416, Dec. 1987.

[8] M. Reyzábal, "Las Competencias Comunicativas y Lingüísticas, Clave para la Calidad Educativa", (online) *REICE,* vol. 10, no. 4, pp. 68, 2012. Available: http://www.rinace.net/reice/numeros/arts/vol10num4/art5.pdf

[9] ABET, "Criteria for Accrediting Engineering Programs, 2017 - 2018" (online) *ABET*, 2016. Available: http://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2017-2018/

[10] B. M. Moskal, "Scoring rubrics: what, when and how?" (online) *Practical Assessment, Research & Evaluation*, vol. 7, no. 3, Sept 2005. Available: http://PAREonline.net/getvn.asp?v=7&n=3

[11] ABET, "Criteria for Accrediting Engineering Programs," (online) *ABET.* Oct 2016. Available: http://www.abet.org/wp-content/uploads/2016/12/E001-17-18-EAC-Criteria-10-29-16-1.pdf

[12] C. A. Mertler, "Designing scoring rubrics for your classroom," (Online) *Practical Assessment, Research & Evaluation,* vol. 7, no. 25, 2001. Available: http://PAREonline.net/getvn.asp?v=7&n=25

[13] B. M. Moskal, "Scoring rubrics: what, when and how?" (online) *Practical Assessment, Research & Evaluation*, vol. 7, no. 3, Sept 2005. Available: http://PAREonline.net/getvn.asp?v=7&n=3

[14] L. Brodie, and P. Gibbings. "Comparison of PBL Assessment Rubrics," *Proceedings of the Research in Engineering Education Symposium,* 2009, Palm Cove, QLD.

[15] ABET, "General Criterion 3. Student Outcomes," (Online) *ABET.* 2017 – 2018. Available: http://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2017-2018/