WIP: A systematic approach for outcome–based assessment

F. Belkhouche
Department of Computer, Electrical and Software Engineering
Shippensburg University of Pennsylvania
Shippensburg, PA, USA
email: fbelkhouche@ship.edu

K. Bendjilali
Department of Mathematics and Computer Science
Raritan Valley Community College
Branchburg, NJ, USA
email: khadidja.bendjilali@raritanval.edu

***Abstract*—This research-to-practice WIP paper describes methods for improving assessment. Assessment is crucial for program accreditation and continuous improvement in engineering education. Outcome–based learning is a recent student–centred approach that can be used perform assessment. Under this paradigm, clear outcomes related to the knowledge and skills that the students should have after completing the course are identified and assessed. The student attainments are measured against the outcomes using clear and well–defined metrics. In this paper, an outcome–based strategy is proposed and used for assessing courses in control systems. In order to improve assessment reliability, the learning outcomes are broken down into sub-outcomes, which are used as performance indicators. The advantage of the proposed method is twofold, first it provides more accurate and reliable metrics for measuring student attainment, and second, the assessment results are personalized, allowing to evaluate individual students and identify means to help them. This approach also allows to identify common mistakes and misconceptions. Data are collected for several semesters and analyzed. It turns out that the proposed approach allows to obtain more accurate and reliable assessment results.**

I. Introduction

Outcome-based learning plays an important role in modern engineering education [1]. This paradigm focuses on what the students are able to do after completing the course or the program. Outcome–based learning is goal-oriented where the desired objectives are set based on well defined criteria. In this case, learning can be seen as a closed loop system, where the difference between the desired outcomes and the measured attainment of the outcomes is minimized. While the concept of outcome-based learning is not new, it has recently received considerable attention due to its importance and relationship with assessment and accountability. Unlike many traditional approaches that are teacher-centered, this approach is student-centered and the curriculum is designed based on the student learning outcomes. There exist several ways to represent the outcomes. This study focuses on course–level outcomes, which are typically area dependent and more specific than program outcomes. For this study, the outcomes focus on skills and knowledge related to controls engineering and belong mostly to the cognitive domain. The students are expected to achieve a desired level of performance after completing the course. The purpose is to graduate engineers who have the knowledge and skills to formulate and solve control engineering problems as professionals.

There are several arguments against outcome-based learning, some authors argue that learning should be broader and not focused on specific outcomes. By defining the outcomes, we put additional restrictions that may limit the content and the coverage depth and therefore the instructors' freedom in designing the curriculum.

Note that outcome–based learning is not restricted to engineering education [2]. It has been proposed as an educational model in many areas including the medical and business fields. Assessment plays an important role in the quality control of any educational program. Another important advantage of outcome based learning is its compatibility with the assessment of the student learning. Outcome–based learning allows to design a comprehensive and systematic approach to assessment, where reliability and validity are the primary goals.

The importance of accurate and reliable measurements is obvious in all feedback systems. Typically, rubrics are used as a means to measure the student attainment of the outcomes. However, different instructors may interpret rubrics differently. Some instructors focus on the completeness of the solutions, other instructors focus on the concept correctness, others may focus on the solution strategy and its implementation without paying too much attention to the final results. In her seminal book [3], Suskie cites a list of biases that affect the assessment process. The list includes ”center tendency errors, Halo effect bias, first impression effect bias, similar-to-me bias”, etc. [3]. The relationship between program or course assessment and outcome based learning is straightforward. Outcome–based learning allows to set the reference and assign rules for the measurement of the student achievement. One of the most important aspects of outcome-based learning is how to measure the student attainment of the outcomes. One of the goals of this paper is to reduce assessment biases. Assessing higher level learning skills using rubrics alone is difficult, even with the most precise and well written rubrics. Therefore, a scale for assessment based on rubrics alone will lack accuracy and reliability. In order to reduce ambiguity and improve reliability and validity, we propose a more modular approach where the assessment tasks are designed to focus on a specific aspect of the learning outcome. Therefore, each outcome is divided into relevant and more specific sub-outcomes that can be easily assessed.

The contribution of this paper consists of the following aspects:

- Develop strategies for accurate and objective measurement of the student attainment of the course learning

outcomes, and reduce variability in assessment.
- Build student portfolios based on their attainment of the outcomes. We refer to this as personalized assessment.

It is important to note that this study does not use any grades or marks in the assessment process. Test scores and grades are not always based on clear and consistent standards, and therefore do not necessarily reflect the student attainment of the learning outcomes [4]. While this paper targets introductory controls courses, the proposed approach is can be easily generalized to other courses and areas, including other engineering disciplines.

II. Course description and outcomes

There exist two approaches for assessing student learning outcomes: direct and indirect [5]. This paper focuses on direct and embedded assessment in control systems. The course under consideration is "Introduction to Feedback Systems". This is a required course for electrical and mechanical engineering majors at most institutions. The course outcomes are designed to reflect the knowledge and skills gained by the students after completing the course. For accurate assessment, it is important to clearly and accurately describe the skills and knowledge gained by students and how attainment is reached. Therefore, establishing appropriate learning outcomes is crucial. The learning outcomes are explicit statements about expectations that must be clear and concise, and as close as possible to complete. If the learning outcomes are not clear, it will be difficult to perform reliable assessment. Embedded assessment of the course activities presents several advantages such as direct measurement of the student attainment. It also allows to identify students who lacked in the attainment of the outcomes and track student performance. Introduction to Feedback Systems is the first course in control systems. The topics covered in this course are standard and traditional topics covered in most control systems/engineering courses across the country. The prerequisite for this course is Signals and Systems course. Students at this stage should have knowledge of basic circuits and calculus III. The course catalog is used to define four learning outcomes that reflect the most important aspects of the course. The learning outcomes are listed below.

A. Learning outcomes

After successfully completing the course, students will be able to:

- **Outcome A:** Use transfer functions and state space models to represent real world control systems.
- **Outcome B:** Analyze transient behavior, steady state errors, and stability of control systems.
- **Outcome C:** Design feedback control systems using transform techniques to achieve desired system specifications.
- **Outcome D:** Use contemporary simulation tools to simulate and solve control engineering problems.

These outcomes focus on the cognitive learning domain.

III. Concept maps

According to the Ausubelian theory of learning, students learn more effectively when they can link concepts with each other, including previously learned concepts. A concept map is a node-edge graphical representation of the concepts and the relationship between them [6]. Concept maps have been used in many areas including nursing [1] and engineering [7], [8]. Typically the concepts are represented by boxes or nodes and the relationships by arrows or edges. There exist several types of concept maps such as the spider and hierarchical maps. Concept maps are generally combination of spider and hierarchical structures with many dependencies and across links.

The content of the course can be seen as a complex interconnected system of concepts. Concept maps are created in several steps [7] as follows:

- Define the key concepts
- Define more specific concepts based on the key or general concepts.
- Establish the relationship between concepts at each level.

Concept maps can be used at each level, including course and program level. Various authors proposed to use them for assessment purpose.

IV. The assessment process, questions and expectations

Adequate assessment should facilitate continuous improvement to achieve the optimal attainment of the learning outcomes. Rubrics are widely used for evaluating student work. Therefore, they need to be designed carefully to avoid confusion, otherwise, the final results may vary depending on the instructor's interpretation of the rubric. Similar to any measurement system, it is important to improve the reliability and validity of the assessment process. One of the objectives of the current study is to determine how well the course outcomes are attained using a new assessment strategy that focuses on the sub-outcomes. In order to diversify the assessment tasks, each outcome is assessed using several sub-outcomes. The assessment process of the learning outcomes should be accurate, objective, and bias-free. In our approach, the assessment tasks are designed to satisfy the following conditions:

- Each learning outcome is divided into a set of sub-outcomes. The sub-outcomes are used as performance indicators
- The examination and assessment questions are aligned with the learning outcomes
- The assessment questions focus on the sub-outcomes
- The sub-outcomes are assessed independently (in isolation). In other words, each question targets a single sub-outcome.
- The math/calculation needed to get the results are minimized (except in the case when the outcome is related to the student ability to use mathematics to solve engineering problems.)

TABLE I
PROPOSED BENCHMARKS FOR OUTCOME ATTAINMENT

Score	Expectation
More than 95%.	exceed expectations
Between 70% and 95%	meet expectations
Less than 70%.	below expectation

- For each question, the assessment result is binary, that is 1 if the student answers the question correctly and 0 otherwise.
- The assessment questions are clear and directly linked to the outcomes.

The purpose of these rules is to minimize any errors that may result from the ambiguity of the rubric or the bias of the assessor.

The attainment and expectations are divided into three levels or categories as summarized in table I. The desired expectation level is set at 70%.

To be practical, assessment techniques need to be sustainable. Assessment activities should be performed continuously during the semester in parallel with the learning activities. Examples of assessment tasks include in-class exams, homework assignments and quizzes. In–class assessment tasks are given a higher weight to achieve more reliable assessment.

The learning outcomes focus on the skills and knowledge related to control engineering. Typically, the recommended number of outcomes is between three and seven. While it is possible to capture the key elements using only a small number of outcomes, it is more difficult to perform objective assessment without clear and measurable performance indicators. Therefore, the outcomes are broken down and represented under the form of modular sub-outcomes that are used as performance indicators.

A. Learning Sub-outcomes

The course outcomes are divided into distinct elements that can be directly measured and assessed in isolation. These sub-outcomes are treated as performance indicators. The list of sub-outcomes is listed below.

- Outcome A:
  - A1: Use the frequency domain representation to model control systems
  - A2: Use the state space representation to model control systems
  - A3: Calculate linearized models of nonlinear systems
  - A4: Convert state space representation to frequency domain
  - A5: Convert frequency domain representation to state space
  - A6: Represent control systems using signal flow graphs
  - A7: Represent control systems using block diagrams
  - A8: Calculate system's response from frequency domain representation
  - A9: Calculate system's response from state space representation
  - A10: Use similarity transformations to find equivalent systems
  - A11: Describe common nonlinearities in control systems
  - A12: List the major components of a closed loop control system
  - A13: Describe the major components of closed loop control systems
- Outcome B:
  - B1: Analyze the transient behavior of control systems.
  - B2: Analyze steady state errors of control systems for different input signals.
  - B3: Analyze stability of control systems in the frequency domain.
  - B4: Analyze stability of control systems in state space.
  - B5: Analyze stability using graphical methods
  - B6: List the common desired requirements in control systems
  - B7: Define the type of the system
- Outcome C:
  - C1: Explain the concept of controllability of control systems
  - C2: Explain the concept of observability of control systems
  - C3: List the advantages of closed loop configuration
  - C4: Design proportional controller
  - C5: Design proportional and integral controller
  - C6: Design lead and lag compensators
  - C7: Explain the effect of controllers and their limitations
  - C8: Design controllers using graphical methods
  - C9: Evaluate the closed loop system characteristics
  - C10: Compare between open loop and closed loop characteristics
  - C11: Write the transfer function of simpler controllers
  - C12: Prioritize the requirements when tradeoffs are present
- Outcome D:
  - D1: Write code to obtain system's response to different inputs.
  - D2: Write code to simulate open and closed loop control systems
  - D3: Use Simulink to simulate control systems
  - D4: Compare between systems and methods using simulation

The total number of sub-outcomes is 36. It is clear that this is not a unique set of sub-outcomes. It is possible to remove or add other sub-outcomes depending on the needs and the programs focus. Bloom's taxonomy provides a unique framework for writing the outcomes. Table II shows a mapping between the learning outcomes and the cognitive domain levels according to Bloom's taxonomy. All six levels in the cognitive domain are covered in this course. Table III shows a breakdown of the sub-outcomes and their relationship with the cognitive domain levels. The highest coverage numbers correspond to outcomes A and C, which are related to system modeling and control design, followed by outcome B and then outcome D. The outcomes as stated previously will cover only three cognitive levels according to Bloom's taxonomy: application, analysis, synthesis. However, at the sub-outcome level, each outcome has several cognitive levels. This is another advantage of the method since it allows for a more accurate description and mapping of each outcome to the cognitive domain learning levels.

V. Alignment between outcomes, curriculum and examination

The lack of alignment between the learning outcomes and the teaching strategy and contents is problematic [3]. In outcome–based learning, the curriculum, the assessment procedures and the examinations are aligned [3]. It is quite common for undergraduate students to study for exams, they tend to focus on the topics that are more likely to be covered in exams. Unfortunately, this happens even to the best students. In general students cannot be blamed for this attitude. In order to address this problem, the assessment process should be decoupled from the students' grades and aligned with the outcomes. In other words, the exams should reflect the most important concepts and skills related to the outcomes of the course for the assessment to be reliable and valid.

VI. Student portfolios

Typically student portfolios refer to collections of student work [3]. Portfolios can vary considerably depending on the purpose. The purpose of portfolios is multi-fold, including gathering information about common misconceptions and developing plans to help under-performing students. In this paper, we propose to construct assessment portfolios from assessment results. Assessment portfolios refer to information about the performance and the attainment of each outcome by the students. The information does not include the grade. Similar to the grades, this information is kept confidential. They can be used for both specific class outcomes attainment and program outcomes attainment. While this paper focuses on a specific controls course, it is possible to include more courses and construct more complete and comprehensive portfolios. This allows to build a baseline data for the students entering the program (freshman and transfer students) and monitor their progress and academic growth. The student assessment portfolios allow to perform personalized assessment and help more students more effectively.

In the proposed approach, student portfolios include the outcome and the cognitive level attainment according to Bloom's taxonomy. In both cases the portfolio includes the percentage of tasks accomplished successfully by the student and the student placement in terms of the expectations. A general format for student portfolio is shown in table VII.

A. Central tendency and variability

Central tendency analysis is widely used to evaluate assessment results. Several measures are used, including the mean, mode, and the median values. While the assumption of normality is made in many studies, our study indicates that the results mostly follow a β distribution (skewed distribution) and are rarely a pure Gaussian. For this distribution, the mean, median and mode are not equal in general. Therefore, they provide different information. Another important measure used to evaluate assessment results is the variability. Variability refers to the difference in the attainment of the learning outcomes. Two measures are used for variability, the range and the variance. The range is simply defined as the difference between the highest and the lowest values in the attainment. In quality control, variability is defined as the reciprocal of quality, therefore, one way to improve quality is by reducing variability.

B. Correlation

Correlation measures can be used to determine the degree of relatedness between the attainment of the different outcomes. The correlation matrix representing the attainment of the outcomes is given by

$$R = \begin{bmatrix} 1.0000 & 0.7835 & 0.8535 & 0.3963 \\ 0.7835 & 1.0000 & 0.7309 & 0.3658 \\ 0.8535 & 0.7309 & 1.0000 & 0.4205 \\ 0.3963 & 0.3658 & 0.4205 & 1.0000 \end{bmatrix} \tag{1}$$

This matrix shows a strong correlation between outcomes A, B and C. Outcome D is less correlated with the others. The smallest correlation is between outcome B and outcome D. Similar results can be seen by examining the number of students who did not meet expectations for each outcome. An example of the percentage of students who did not meet the expectation in one or two outcomes is shown in table V. The diagonal elements represent the percentage of students who did not meet the expectation is a specific outcome. The other elements represent the percentage of students who failed to meet the expectations in two outcomes that correspond to the row and the column.

Another way to represent correlation is by finding the percentage of students who failed one or more outcomes.

C. Internal consistency reliability analysis

Internal consistency reliability is a quality index of the assessment data. There exist several indices that can be used to gauge reliability. Typically, these indices take values between 0 and 1. A high score indicates better quality measurement. One of the objectives of the proposed approach is to improve the reliability of the measurements by using the sub-outcomes as a tool to achieve repeatability. While there exist several ways to define reliability in assessment, we are particularly interested in the internal consistency reliability, since it is the most relevant definition to this study. For each outcome, the total observed assessment score is the sum of the true assessment score and the errors in measurement. The errors have different sources including

- Sources associated with the assessment task such as short tests and assessment questions that are ambiguous or unclear.
- Sources related to the test administration or environment.

Learning Outcomes				
Cognitive domain levels	Outcome A	Outcome B	Outcome C	Outcome D
Knowledge	✓	✓	✓	
Understanding	✓	✓	✓	
Application		✓		✓
Analysis	✓		✓	
Synthesis			✓	
Evaluation			✓	✓
Total	36.11 %	19.44 %	33.33 %	11.11%

TABLE II

LEARNING OUTCOMES AND THEIR RELATIONSHIP WITH THE COGNITIVE DOMAIN LEVELS

Cognitive domain levels	Sub-outcomes covered	Percentage
Knowledge	A12, B7, C3, C11	11.11 %
Understanding	A6, A7, A11, A13, B6, C1, C2, C7	22.22 %
Application	A1, A2, A4, A5, A10, D1, D2, D3	22.22%
Analysis	A3, A8, A9, B1, B2, B3, B4, B5	22.22 %
Synthesis	C4, C5, C6, C8	11.13 %
Evaluation	C9, C10, C12, D4	11.11 %

TABLE III

SUB-OUTCOMES AND THEIR RELATIONSHIP WITH THE COGNITIVE DOMAIN LEVELS

Learning Outcomes and cognitive domain levels	Percentage of tasks accomplished successfully	Below expectations	Meet expectations	Exceed expectations
Outcome A	x_1%	✓		
Outcome B	x_2%	✓		
Outcome C	x_3%	✓		
Outcome D	x_4%		✓	
Total	100%			
Knowledge	y_1%		✓	
Understanding	y_2%	✓		
Application	y_3%	✓		
Analysis	y_4%		✓	
Synthesis	y_5%		✓	
Evaluation	y_6%	✓		
Total	100%			

TABLE IV

AN EXAMPLE OF ASSESSMENT PORTFOLIO

Outcome	A	B	C	D
A	0.2632	0.1404	0.1930	0.0702
B	—	0.1579	0.1228	0.0526
C	—	—	0.2456	0.0877
D	—	—	—	0.1754

TABLE V

PERCENTAGE OF STUDENTS WHO FAIL TO MEET EXPECTATIONS FOR ONE OUTCOME OR TWO OUTCOMES

- Sources associated with the test takers, such as the mental state at the time of the test, which includes stress, fatigue, etc.
- Sources associated with the assessor or the scoring process. These include imprecise rubrics and assessor's bias.

A highly reliable assessment result implies that the changes in the measurement are small if the assessment tasks were repeated more times. One of the most widely used indices for internal consistency reliability is the Cronbach's coefficient α, which is defined as follows [9], [10]:

$$\alpha_o = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K}\sigma_i^2}{\sigma_o^2}\right) \quad (2)$$

In general, K represents the number of test items. In the proposed formulation, it represents the number assessment tasks used for a particular outcome, σ_o^2 is the variance of outcome o and σ_i^2 is the variance of assessment task i. An example illustrating coefficient α for each outcome is shown in table VI. Other indices such as Kuder-Richardson [10] and split half index can be used.

Outcome	A	B	C	D
Cronbach's coefficient α	0.92	0.71	0.82	0.59

TABLE VI

COEFFICIENT α FOR EACH OUTCOME

A general rule of thumb for coefficient α is summarized in table VII. This table indicates that the internal consistency is excellent for outcome A, good for outcome C, acceptable for outcome B, and poor for outcome D. While the reliability of the proposed method is good for most outcomes, outcome D has the minimum acceptable reliability and therefore more attention is needed for this outcome. Improvement can be

Cronbach's coefficient α	Internal consistency
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$\alpha < 0.5$	Unacceptable

TABLE VII
CRONBACH'S COEFFICIENT α FOR RELIABILITY [10]

achieved by adding more sub-outcomes or more assessment tasks in outcome D.

D. Outcome agreement and disagreement indices

Another way to evaluate the relationship or cause effect between the outcomes is by using the indices of agreement/disagreement. These indices are defined as follows

- Agreement index:

$$A_{o_1,o_2} = a + d \tag{3}$$

- Disagreement index

$$D_{o_1,o_2} = b + c = 1 - A_{o_1,o_2} \tag{4}$$

where

- a is the percentage of students who met the expectations in outcomes o_1 and o_2
- b is the percentage of students who met the expectations in outcomes o_1 but not o_2
- c is the percentage of students who met the expectations in outcomes o_2 but not o_1
- d is the percentage of students who did not meet the expectations in neither outcome.

It is possible to generalize the concept of agreement/disagreement index to include more than two outcomes. One particular case is to consider the agreement index for all outcomes. Note that a higher value of the agreement index may indicate a lack of orthogonality between outcomes.

E. Standard error of measurement (SEM)

Another index that can be used to gauge the assessment results is the standard error of measurement defined as follows

$$SEM_o = \sigma_o \sqrt{1 - \alpha_o} \tag{5}$$

where σ_o is the standard deviation of the assessment results for outcome o. The standard error of measurement depends on the reliability of the measurements and the standard deviation. It is close to zero for highly consistent results, and equal to the standard deviation when the measurements are not consistent.

VII. CONCLUSION

Assessment is necessary for continuous improvement as it allows to perform accurate diagnosis of the problems associated with the learning/teaching process. This paper proposes simple ways to improve the assessment reliability while ensuring validity at the same time. In the first step, clear outcomes that reflect the most important aspects of the course are proposed. These outcomes are then divided into sub-outcomes that are assessed independently. The sub-outcomes allow to transfer the problem from a single item scale to a multi–item scale. While the proposed approach does not suggest to completely eliminate rubrics, it aims at minimizing the use of rubrics since they can lead to subjective interpretations of the student attainment of specific outcomes. The sub-outcomes provide valuable ways to evaluate the internal consistency reliability of the assessment results and construct representative student portfolios.

REFERENCES

[1] B. J. Daley, S. Morgan, and S. B. Beman, "Concept maps in nursing education: A historical literature review and research directions," *Journal of Nursing Education*, vol. 55, no. 11, pp. 631–639, 2016.

[2] M. H. Davis, "Outcome-based education," *Journal of veterinary medical education*, vol. 30, no. 3, pp. 258–263, 2017.

[3] L. Suskie, *Assessing Student Learning: A Common Sense Guide*. Jossey-Bass, 2009.

[4] T. El-Maaddawy and C. Deneen, "Outcomes-based assessment and learning: Trialling change in a postgraduate civil engineering course," *Journal of University Teaching and Learning*, vol. 14, no. 1, p. 10, 2017.

[5] G. B. Wright, "Student-centered learning in higher education," *International Journal of Teaching and Learning in Higher Education*, vol. 23, no. 1, pp. 92–97, 2011.

[6] M. Buldua and N. Buldub, "Concept mapping as a formative assessment in college classrooms: Measuring usefulness and student satisfaction," *Social and Behavioral Sciences*, vol. 2, p. 2099–2104, 2010.

[7] J. Turns, C. J. Atman, and R. Adams, "Concept maps for engineering education: A cognitively motivated tool supporting varied assessment functions," *IEEE Transactions on Education*, vol. 43, no. 2, pp. 164–173, 2000.

[8] M. Kilic and M. Cakmak, "Concept maps as a tool for meaningful learning and teaching in chemistry education," *International Journal on New Trends in Education and Their Implications*, vol. 4, pp. 152–164, 2013.

[9] R. A. Peterson, "A meta-analysis of cronbach's coefficient alpha," *Journal of Consumer Research*, vol. 21, no. 2, pp. 381–391, 1994.

[10] M. B. Miller, "Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 2, no. 3, pp. 255–273, 1995.