Digitally Automated Assessment of Outcomes Classified Per Bloom’s Three Domains and Based on Frequency and Type of Assessments

Wajid Hussain & Mohammad F. Addas
Islamic University in Madinah

April 2016

One of the primary reasons outcomes information is not utilized for Continuous Quality Improvement (CQI) is that the information collected is insufficient to make improvement decisions due to impractical manual processes that are either too exhaustive to complete for timely measurement and reporting, or too minimal for basic fulfillment of accreditation requirements. Massive amounts of outcomes data collected from various stages of curriculum delivery is a critical requirement for informing improvement decisions. Therefore, manual assessment, documentation and reporting systems are major factors that exacerbate the implementation of streamlining activities which are necessary to integrate improvement efforts of several stakeholders in an academic CQI process. In an age of technological advancement, use of digital technology allows for the collection of various evidence sources. The Faculty of Engineering at the Islamic University outlined five crucial elements of their outcomes assessment methodology which fully supports automation and digital technology based assessment/documentation/reporting systems to collect, analyze and utilize outcomes data to establish meaningful CQI and not just fulfill accreditation requirements.

1. MEASUREMENT OF OUTCOMES IN ALL COURSE LEVELS OF A PROGRAM CURRICULUM (refer Figure 1).

Generally institutions classify courses of a program curriculum into three levels: introductory, reinforced and mastery with outcomes assessment data measured for the mastery level courses in order to streamline the documentation and effort needed for an effective program evaluation. This approach presents a
major deficiency for CQI in a student centered outcomes-based education model since performance information of a graduating batch of students collected at just the mastery level to measure program Student Outcomes (SOs) is at a final phase of a typical quality cycle and too late for implementation of remedial efforts for performance failures of the students in consideration. A holistic approach for a CQI model would require a systematic measurement of performance indicators in all three of Bloom’s domains of learning and their corresponding categories of learning levels for all course levels of a program’s curriculum.

![Figure 1: Multiple course levels and Pls classified per Bloom’s 3 domains learning levels utilized for outcomes measurement**](image)

2. FACULTY COURSE ASSESSMENT REPORT (FCAR) UTILIZING THE EAMU PERFORMANCE VECTOR METHODOLOGY

EvalTools® 6 is chosen as the platform for outcomes assessment since it employs the unique Faculty Course Assessment Report (FCAR) and EAMU performance vector methodology (J. Estell, J. Yoder, B. Morrison, F. Mak, 2012) which facilitate the use of existing curricular grade giving assessments for outcomes measurement and help in achieving a high level of automation of the data collection process (Figure 2.), feature-rich pick-and-choose assessment/reporting tools, and the flexibility to provide customized features (www.makteam.com, 2015).

The EvalTools® 6 FCAR module provides summative/formative options and consists of the following components: course description, COs indirect assessment, grade distribution, COs direct assessment, assignment list, course
reflections, old action items, new action items, student outcomes assessment and performance indicators assessment.

Figure 2: Comparative study of the advantages of automation in outcomes assessment achieved with EvalTools® 6 + FCAR + EAMU versus other tools © 2015 Wajid Hussain

![Figure 2: Comparative study of the advantages of automation in outcomes assessment achieved with EvalTools® 6 + FCAR + EAMU versus other tools © 2015 Wajid Hussain](image)

Table: Specification of EAMU performance indicator levels:

<table>
<thead>
<tr>
<th>Category</th>
<th>General Description</th>
<th>Letter Grade</th>
<th>Nominal Indicator Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>Student applies knowledge with virtually no conceptual or procedural errors</td>
<td>E</td>
<td>100%</td>
</tr>
<tr>
<td>Adequate</td>
<td>Student applies knowledge with no significant conceptual errors and only minor procedural errors</td>
<td>A</td>
<td>75.0% - 90.0%</td>
</tr>
<tr>
<td>Minimal</td>
<td>Student applies knowledge with occasional conceptual errors and only minor procedural errors</td>
<td>M</td>
<td>50.0% - 75.0%</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>Student makes significant conceptual and/or procedural errors when applying knowledge</td>
<td>U</td>
<td>0.0% - 50.0%</td>
</tr>
</tbody>
</table>

Table: Heuristic rules for performance vector tables (PVT):

<table>
<thead>
<tr>
<th>Category</th>
<th>General Description</th>
<th>Scale (out of 5)</th>
<th>Maximum Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Flag</td>
<td>Any performance vector with an average below the defined scale and a level of unsatisfactory performance that exceeds defined percentage in the U column</td>
<td>below 3.3</td>
<td>&gt; 10%</td>
</tr>
<tr>
<td>Yellow Flag</td>
<td>Any performance vector with an average below the defined scale or a level of unsatisfactory performance (U) that exceeds the defined percentage, but not both</td>
<td>below 3.3 or &gt; 10%</td>
<td></td>
</tr>
<tr>
<td>Green Flag</td>
<td>Any performance vector with an average that is at least greater than the defined scale and no indication of unsatisfactory performance (U)</td>
<td>&gt;= 4.0</td>
<td></td>
</tr>
<tr>
<td>No Flag</td>
<td>Any performance vector that does not fall into one of the above categories</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3: Performance criteria: EAMU PI levels and heuristic rules for Performance Vector Tables (PVT) adopted by the Faculty of Engineering at the Islamic University of Madinah

![Figure 3: Performance criteria: EAMU PI levels and heuristic rules for Performance Vector Tables (PVT) adopted by the Faculty of Engineering at the Islamic University of Madinah](image)
The FCAR uses the performance vector, conceptually based on a performance assessment scoring rubric developed by Miller and Olds (R. L. Miller, B. M. Olds, 1999) to categorize aggregate student performance.

The EAMU performance vector (Figure 3) counts the number of students that passed the course whose proficiency for that outcome was rated Excellent, Adequate, Minimal, or Unsatisfactory. Program faculty report failing course outcomes (COs), ABET student outcomes (SOs), performance indicators (PIs), comments on student indirect assessments and other general issues of concern in the respective course reflections section of the FCAR. Based upon these course reflections, new action items are generated by the faculty. Old action items status details are carried over into the current FCAR from the information generated during the previous offering for this specific course. Modifications and proposals to a course are made with consideration of the status of the old action items (W. Hussain, M.F. Addas, 2015).

3. DIGITAL DATABASE OF SPECIFIC PERFORMANCE INDICATORS (PIs) CLASSIFIED PER BLOOM’S REVISED 3 DOMAINS OF LEARNING AND THEIR ASSOCIATED LEVELS (according to the 3-Level Skills Grouping Methodology) (W. Hussain, M. F. Addas and Mak F., ASEE 2016)

An important observation made by the Faculty of Engineering is that Bloom’s 3 learning domains present an easier classification of specific PIs for realistic outcomes assessment versus other models that categorize learning domains as knowledge, cognitive, interpersonal, communication/ IT/numerical and/or psychomotor skills. In addition, categories of learning domains which seem very relevant for the engineering industry and career-related requirements may not be practically easy to implement when it comes to classification, measurement of PIs, and realistic final results for CQI measurement.

A hypothetical Learning Domains Wheel as shown in Figure 4 was developed by the Faculty of Engineering to analyze the popular learning domains models available, including Bloom’s, with a perspective of realistic measurement of outcomes based on valid PIs classification that does not result in a vague indicator mechanism for CQI in engineering education. Learning domains categories mentioned in this paper specifically refer to broad categories with well-defined learning levels selected for the classification of specific PIs. The Learning Domains Wheel was implemented with Venn diagrams to represent details of the relationship of popular learning domains categories, interpersonal skills, and the types of knowledge.

The cognitive domain involves acquiring factual, conceptual knowledge dealing with remembering facts and understanding core concepts. Procedural and metacognitive knowledge deal essentially with problem solving, which includes problem identification, critical thinking and metacognitive reflection. Remembering facts, understanding concepts and problem solving are essential, core and universal cognitive skills that would apply to all learning domains. Problem identification, definition, critical thinking and metacognitive reflection are some of the main elements of problem solving skills. These main elements of problem solving skills apply to all levels of learning for the three domains. Activities related
to any learning domain require operational levels of four kinds of knowledge: factual, conceptual, procedural and metacognitive that are proportional to the expected degree of proficiency of skills for proper completion of tasks. For example, successfully completing psychomotor tasks for solving problems involves acquiring very specialized proportions of factual, conceptual, procedural and metacognitive knowledge of various physical processes with accepted levels of their activities skills proficiency. Similarly, an affective learning domain activity, such as implementing a code of professional ethics, involves acquiring factual, conceptual, procedural and metacognitive knowledge related to industry standards, process of application, level of personal responsibility and impact on stakeholders. Hence, the psychomotor and affective domains skills overlap with the cognitive domain for the necessary factual, conceptual, procedural and metacognitive areas of knowledge.

Figure 4: The Learning Domains Wheel for snapshot analysis and selection of learning domains categories to achieve realistic outcomes measurement with easier PIs classification process © 2015 Wajid Hussain
The learning domains categories such as interpersonal, IT, knowledge, cognitive, communication, numerical skills etc., exhibit significant areas of overlap as shown in the Learning Domains Wheel in Figure 4. This large overlap of skills within multiple learning domains presents a serious dilemma to engineering programs in the PIs classification and measurement process. A difficult choice must be made whether to select the most appropriate learning domain category and discard the others or repeat mapping similar PIs to multiple learning domain categories for each classification. Defining the learning levels for the overlapping categories to precisely classify PIs would also be challenging. Finally, learning domain categories with significant areas of overlap would result in the repeated measurement of common PIs in multiple domains and the accumulation of too many types of PIs in any single learning domain category, thus obscuring specific measured information. Therefore, for practical reasons the categories of learning domains have to be meticulously selected with a primary goal of implementing a viable PIs classification process to achieve realistic outcomes measurement for program evaluation.

Crucial guidelines were logically derived from the Learning Domains Wheel for the selection of the learning domains categories as follows:

1. Very broad learning domains categories consist of many skills sets that will present difficulty in the classification of PIs when grouped with other categories and will result in the redundancy of outcomes data; for example, interpersonal skills grouped with IT, communication or psychomotor, etc.
2. Avoid selection of any two skills sets as learning domains categories when one is an absolute subset of another. Just select either the most relevant one or the one which is a whole set. For example, select cognitive or numeric skills, but not both; if both are required, select cognitive as a category since it is a whole set. Numeric skills, its subset, can be classified as a cognitive skill.
3. If selecting a certain skills set that is a whole set as a learning domains category, then it should not contain any other skills sets which are required to be used as learning domains categories; e.g., do not select affective as a learning domains category since it is a whole set if you also plan on selecting teamwork skills as a category.
4. A learning domain category could contain skills sets which will not be utilized for PIs classification; e.g., affective learning domain category containing leadership, teamwork and professional ethics skills sets; leadership, teamwork and professional ethics will NOT be a learning domain category but will be classified as affective domain skill sets.

Bloom’s 3 domains, cognitive, affective and psychomotor, are not absolute subsets of one another. They contain skills sets as prescribed by the 11 EAC ABET SOs which are not learning domains categories. Therefore Bloom’s 3 learning domains satisfy selection guidelines derived from the Learning Domains Wheel and facilitate a relatively easier classification process for specific PIs. Calculation of term-wide weighted average values for ABET SOs using this classification of specific PIs resulted in realistic outcomes data since most of the PIs were uniquely mapped to each of the 3 domains with minimal overlap and redundancy.
Figure 5 shows the design flow for the creation of holistic learning outcomes and their performance indicators for all courses corresponding to introductory, reinforced and mastery levels spanning the curriculum. The Faculty of Engineering studied past research, which grouped Bloom’s learning levels in each domain based on their relation to the various teaching and learning strategies. With some adjustments, a new 3-Level Skills Grouping Methodology was developed for each learning domain with a focus on grouping activities which are closely associated to a similar degree of skills complexity. Figure 6 exhibits this new grouping.
Performance indicators should be specific to collect precise learning outcomes information related to various course topics and phases of a curriculum, while addressing various levels of proficiency of a measured skill. Design of COs and their PIs was meticulously completed by using appropriate action verbs and subject content, thus rendering the COs, their associated PIs, and assessments at a specific skill level—elementary, intermediate, or advanced. Figure 7 shows an example from a civil engineering course. In this example, CO_2: Describe the composition of soil and solve volume-mass relationship equations for soils; and its associated specific PI_5_34: Determine the physical properties of soil using given parameters; measured by assessment Mid Term Q9 are of similar complexity and at the same level of learning. The corresponding category of learning is intermediate-cognitive-applying. Therefore COs would be measured by PIs and assessments strictly following the 3-Level Skills Grouping Methodology.

![Figure 7: Example of a civil engineering course showing CO_2, PI_5_34 and assessment Mid Term Q9 assigned to intermediate-cognitive-applying skill level based on the 3-Level Skills Grouping Methodology]

Ideally, all courses should measure the elementary, intermediate and advanced level skills with their COs, specific PIs and associated assessments. However, introductory level courses should measure a greater proportion of the elementary level skills with their COs, PIs and assessments. On the other hand, mastery level courses should measure more of the advanced, but fewer intermediate and elementary level skills. Figure 8 indicates an ideal learning level distribution of COs and PIs for the introductory, intermediate and mastery level courses.

The measurement of outcomes and PIs designed following such an ideal distribution will result in a comprehensive database of learning outcome information, which will facilitate a thorough analysis of each phase of the learning process and a comparatively easier mechanism for early detection of the root cause of student performance failures at any stage of a student’s education.
The measurement of outcomes and PIs designed following such an ideal distribution will result in a comprehensive database of learning outcome information, which will facilitate a thorough analysis of each phase of the learning process and a comparatively easier mechanism for early detection of the root cause of student performance failures at any stage of a student’s education.

4. SCIENTIFIC CONSTRUCTIVE ALIGNMENT AND UNIQUE ASSESSMENTS TO OBTAIN REALISTIC OUTCOMES DATA (one specific PI per assessment)

Designing any assessment related to specific course content would require considering measurement of the most appropriate performance criteria. For scientific constructive alignment, as opposed to conventional constructive alignment, the contribution of various performance criteria to the total score of an assessment would be defined during assessment design. The performance criteria of interest to be measured by a specific assessment would be given a nearly 70% or more share in the total score distribution and the effect of grading results of the other performance criteria on the total score would be thus rendered negligible. Figure 9 shows an example where a sample unique assessment (quiz 2) with high relative coverage (Q2 7 points) is designed with maximum coverage (70%) of a specific PI_5_12 mapping to a CO3, ABET SO5.

Such assessments or set of questions are said to be unique since they are just used once for measurement of a certain PI. This methodology of implementing unique assessments with high relative coverage of PIs mapping to COs and ABET SOs would ensure realistic measurement of outcomes assessment data for comprehensive continuous improvement.
Relevant assignments termed as "key assignments" are used as assessments for measuring specific PIs related to SOs in each course. Most assessments in courses were formative in application (utilizing the formative option in EvalTools® 6) resulting in an adjustment of teaching and learning strategies by faculty. Since assessments are equivalent to learning in the OBE model it was decided to consider the type of assessments, their frequency of implementation and the learning level of measured specific PIs in Bloom’s 3 domains for course and overall program evaluations. At the course level the types of assessments are classified using the course formats chart to calculate their weighting factors (W. Hussain, M.F. Addas, 2015) which are then applied using the setup course portfolio module of EvalTools® 6. The results are available for view in the FCAR and are used for course evaluations.

The program level SO evaluations employ a weighting scheme which considers the frequency of assessments implemented in various courses for a given term.
to measure PIs associated with specific learning levels of Bloom's domains (W. Hussain et al., ASEE 2016). Figure 10 shows the EE program term 361 composite (cognitive, affective and psychomotor) learning domains evaluation data for 11 ABET SOs. For each SO the counts and aggregate average values of assessments implemented in various courses for measuring PIs associated with the specific learning levels are shown. (Mastery level courses were not offered in term 361).

![Counts and Values of Assessments Implemented for Different PIs in Multiple Courses for SO 1](image)

**Figure 10: EE program term 361 Learning domains evaluations**
Figure 11: Course level CQI with alignment of assessments, teaching & learning strategies according to Bloom’s 3 domains and 3-Skills Levels Methodology**

Figure 11 shows the course level alignment of assessments, teaching & learning strategies to cover the deficiency in measurement of elementary skills thereby...
rendering the assessments formative. (W. Hussain, M.F. Addas, Mak F., FIE 2016). Figure 12 shows program term reviews (SO/PI evaluations) report sample exhibiting CQI efforts, action items, discussions etc. (W. Hussain et al., FIE 2016).

Figure 12: Program term reviews (SO/PI evaluations) report sample exhibiting CQI efforts, action items, discussions etc.

6. ELECTRONIC INTEGRATION OF ADMINISTRATIVE ASSISTANT SYSTEM (AAS), LEARNING MANAGEMENT SYSTEM (LMS) WITH OUTCOMES ASSESSMENT SYSTEM (OAS) AND CONTINUOUS IMPROVEMENT MANAGEMENT SYSTEM (CIMS) FACILITATING FACULTY INVOLVEMENT FOR REALISTIC CQI

7. ELECTRONIC INTEGRATION OF ACTION ITEMS (AIs) GENERATED FROM PROGRAM OUTCOMES TERM REVIEWS WITH STANDING COMMITTEES MEETINGS, TASKS LISTS AND OVERALL CQI PROCESSES (CIMS FEATURE) (W. Hussain et al., ASEE 2016)

A minority of faculty members were initially reluctant to implement digital technology incorporating FCAR methodology and PIs classification per Bloom’s 3 domains. One of the reasons for this resistance was the lack of comprehension of ABET accreditation, latest outcomes assessment processes, and experience regarding their management. Detailed training sessions followed up with extensive technical and intellectual support from the Office of Quality and Accreditation.
for the Faculty of Engineering significantly alleviated their reservations. Various program level sessions held for the development and classification of specific PIs actually galvanized the interest levels of faculty members by providing them with a first-hand learning experience to develop measurable learning outcomes, their PIs and assessments as per Bloom’s 3 domains, and their learning levels. The most difficult aspect of continuous improvement and accreditation efforts for faculty members was to create action items for improvement based upon deficient outcomes assessment data, assign them to the concerned parties or individuals, and follow up for closing the loop. Implementing physical systems to maintain huge amounts of paper-based documentation and manual processes to access specific, on-time information for CQI activity related to closing the loop were specifically the biggest challenges faced by the faculty members.

The Continuous Improvement Management System (CIMS) provided our faculty with efficient streamlining mechanisms for quality improvement efforts by employing very high levels of automation and paper-free digital documentation. Instant electronic access to digital records of single or multi-term outcomes assessment information from program reviews and detailed meeting minutes, action items status of 17 standing committees, essential for CQI efforts, were compelling reasons for an eventual, almost 100% faculty buy-in of the implemented digital systems and outcomes assessment methodologies.

With a majority of positive aspects, one limitation of our system, the allocation of resources to scan paper documents, is currently performed by either the lecturers or teaching assistants. Work is currently in progress to develop state-of-the-art digital systems that automate outcomes assessment development and scoring processes. This technology would integrate with existing digital systems to significantly reduce the overhead related to overall time spent by faculty in the outcomes assessment process and scanning work done by lecturers. In conclusion, we have achieved our goal to evaluate engineering programs based on the automated measurement of PIs classified into the cognitive, affective and psychomotor learning domains of the revised Bloom’s taxonomy.

* Reprinted by permission of Makteam Inc.

** Reprinted by permission of faculty of engineering, Islamic University, Madina, KSA (generated by EvalTools® 6)

*** Reprinted by permission of faculty of engineering, Islamic University, Madina, KSA

† Islamic University of Madinah semester naming system, where first two digits ‘36’ refer to the local year code and the last digit refers to the semester, 1: fall, 2: spring and 3: summer.
REFERENCES


SOURCES OF FURTHER INFORMATION


Information on EvalTools® available at http://www.makteam.com

W. Hussain: Digital Technology for Outcomes Assessment in Higher Education, https://www.youtube.com/watch?v=JaQ0trgk6YE

W. Hussain: Automated Engineering Program Outcomes, Bloom’s Learning Domains Evaluations https://www.youtube.com/watch?v=HAGaoRUrJIE
About NILOA

- The National Institute for Learning Outcomes Assessment (NILOA) was established in December 2008, and is co-located at the University of Illinois and Indiana University.
- The NILOA website contains free assessment resources and can be found at http://www.learningoutcomesassessment.org.
- The NILOA research team has scanned institutional websites, surveyed chief academic officers, and commissioned a series of occasional papers.

NILOA Staff

Natasha Jankowski, Director
Gianina Baker, Assistant Director
Katie Schultz, Project Manager
Erick Montenegro, Communications Coordinator & Research Analyst
Verna F. Orr, Post-Doctoral Researcher

NILOA Senior Scholars

Peter Ewell, Senior Scholar
Pat Hutchings, Senior Scholar
Stanley Ikenberry, President Emeritus, University of Illinois
Jillian Kinzie, Senior Scholar
George Kuh, Founding Director, Senior Scholar
Paul Lingenfelter, Senior Scholar
David Marshall, Senior Scholar

NILOA Sponsors

Lumina Foundation for Education
University of Illinois, College of Education