Effective Assessment of Computer Science Capstone Projects and Student Outcomes

https://doi.org/10.3991/ijep.v10i2.11855

Fatima K. Abu Salem American University of Beirut, Beirut, Lebanon

Issam W. Damaj (⊠) Beirut Arab University, Debbieh, Lebanon i.damaj@bau.edu.lb

Lama A. Hamandi American University of Beirut, Beirut, Lebanon

Rached N. Zantout Rafic Hariri University, Mechref, Lebanon

Abstract—A capstone project is a culminating experience that entails creativity, critical thinking, and advanced problem-solving skills. To that end, capstone projects enable students to prove their abilities, demonstrate their attained skills, and carry out a significant project in their field. In Computer Science Bachelor programs, there is a strong mapping between learning outcomes of capstone projects and all student learning outcomes. This paper presents an assessment framework for capstone courses that allows for sound evaluations of the performance of students and project qualities; besides assessing the student outcomes of the program. The developed framework comprises criteria, indicators, extensive analytic rubrics, and an aggregate statistical formulation. The presented course and framework are supported by the results, analysis, and evaluation of a pilot study for a single institution to explore the effectiveness of the proposed tool.

Keywords—Assessment, Capstone Projects, Computer Science, Accreditation, ABET.

1 Introduction

Capstone projects are a rich resource for assessing the level of attainment of student outcomes and in most of the knowledge acquired by students during their study. This assessment is done towards the end of students' studies giving a true indication of the level of attainment of learning outcomes. However, capstone project assessment is very challenging. It is different from assessment in a regular course. Capstone projects do not feature traditional modes of assessment that include lab assignments, home assignments, period quizzes and term exams. In contrast also, students usually are undertaking learning tasks that are more research oriented rather than application oriented. As a result, evaluators would be employing different evaluation strategies inspired by the wide spectrum reflecting the versatility in their expertise. Another challenge is that a capstone project covers all student outcomes.

It is very important to develop a capstone project assessment tool that will address the above challenges. One should base this tool on performance criteria that adhere to the capstone project, particularly its learning outcomes. It should also be flexible to be used for assessing projects of different nature. Moreover, it should be easy enough to be used by evaluators having different expertise. Indeed, the tool should be clear so that students would be able to prepare themselves accordingly.

In this article, we develop a framework to assess the learning outcomes of capstone projects (CPs), and hence the mapping that binds them to Student Outcomes (SOs) at the program level. Our focal point of interest is in Computer Science and related Computing programs [1]. Our tool follows the ABET Accreditation criteria where we adopt a large part of its terminology. In line with the works of [2] and [3], our tool consists of a suite of criteria and their indicators, supported by thorough rubrics, and concluding with a summative statistical aggregation.

Our manuscript is organized so that in Section 2 we present related work. In Section 3, we develop the measurement scheme and address aspects of assessment criteria as well as indicators. Section 4 presents a discussion on the measurement rubrics, together along the statistical formulation for both evaluating student performance, as well as the attainment of SOs. In Section 5, we delve into the pilot study and present its analysis and evaluation. Section 5 presents the conclusion and offers insight for future work.

2 Related Work

Several investigations are reported in the literature to present effective Capstone Project (CP) course setups, structures, and assessment frameworks in computer science and engineering programs [1-12]. In [1], the author describes a course structure that includes a carefully designed prerequisite course on project management and scheduled milestones. Furthermore, assessment is done using a holistic rubric that enables the measurement of a set of Course Learning Outcomes (CLOs) that are derived from ABET Student Outcomes (SOs). Yousafzai et al. [2] and Damaj et al. [3] present a unified framework which allows for sound evaluations of student performance and CP qualities in addition to assessing SOs. Along the lines of those works, the authors develop a framework consisting of a suite of criteria and their indicators, supported by thorough analytic rubrics, and concluding with a summative statistical formulation. The authors conduct a pilot study examining several capstone projects. The results reveal that after raters' calibration, only a minor difference in average scores exist. In fact, the authors examine the weighted average assessment scores pre and post calibration and report it to be at 3.2%.

Olarte et al. [4] present a study that compares student and staff perceptions of key aspects of CPs, such as, characteristics of projects, competencies of students, in-

volvement of advisor, and perceived learning of students. The study employs a holistic Likert scale of 1–4 to aid the evaluation. Three questionnaires were designed for the student, the advisor, and the committee. The questionnaires are divided into four blocks, one for each of the key elements. The blocks examine project characteristics, student competency, advisor involvement, and the level of learning perceived by students. primary finding of this study is that the students' expectations differ greatly from those of staff.

Furthermore, a supervision typology in Computer Science Engineering CPs is presented in [5]. The study develops and validates an instrument, and then utilizes it to determine different styles of supervision. A questionnaire is developed to survey capstone project advisors at a university during the past two years. A total of 109 surveys are successfully collected. A combination of multivariate statistical methods, such as factorial analysis, is employed. This study distinguished seven main supervision factors: technology, arrangements, keep alive, execution, meetings, management, and reports.

Assessment rubrics were presented for software-related undergraduate capstone projects in [6]. In addition, it was recommended that capstone projects should always undergo a continuous improvement process. A survey was carried out at different Pakistani universities. The survey results highlighted several challenges such as poor process visibility, difference in support and information documents, limited guidelines for project assessment, lack of adequate requirements on software process and documentation, and limited incentives for supervisors. The proposed rubrics specify the key assessment criteria to be assessed using quality levels.

Instead of having a capstone project course, a series of courses include capstone projects is presented by Bihari et al. [7]. Among those is a course in Software Engineering in which control was inverted so that the industrial sponsor has more control and management duties over the project than the faculty member supervising the students. The course was scaled up successfully through developing unique assessment and evaluation tools to monitor, measure and fairly assess a wide spectrum of projects. Students are evaluated based on several presentations done at various points in the quarter, in-class weekly reports, industrial sponsor feedback, project workbook, poster and an individual report. The rubrics are designed to be immune to variability in projects, variability in industry sponsors and variability in technology. Each deliverable is evaluated on a combination of the choice of the method as well as the execution of the method.

Moreover, the meetings-flow (MF) method is evaluated in, in [8], terms of its effect on teams undergoing their capstone projects. Previous studies have shown that MF is beneficial for monitoring student work and product quality. In this study, it was empirically proven that MF enhances a team's communication and coordination and creates a balance between the contributions of all team members. However, MF was observed to have small influence on team cohesion.

In addition, an instrument is developed, by Pérez et al [9], to determine the different styles of supervision for advisors of capstone projects. Six supervision styles were identified based on seven factors of supervision. The identified supervision styles and

factors can be used to guide advisors on how to supervise students as well as where to assess competencies of students and how to give meaningful feedback to the students.

3 The Measurement Scheme

3.1 Background

According to [2, 3], the CP's learning outcomes can be associated with or replaced completely by the SOs at the program level. According to ABET, student learning outcomes capture the knowledge that students are to attain or to perform by graduation time. There are six student learning outcomes associated with the Computer Science program and as set out by ABET [10]. The CP is the most critical juncture in a student's undergraduate journey where they get exposed to a significantly practical experience. The CP is the very first encounter in their educational timeline that propels students towards an area of their own interest. With contemporary worldwide societal challenges, it also becomes imperative to train students on using their knowledge for the service of public good. As such, the CP provides an optimal venue where students can aim to develop technologies and analytics for addressing societal challenges.

As discussed earlier, our assessment approach is dedicated for Computer Science programs and related Computing fields. In our approach we consider a CP course taking place in one semester only. The CP course aims to enhance students' skills with practical experience giving them the opportunity to integrate knowledge accumulated in different courses. The pre-requisites for the CP are three junior and senior level courses following a three-year undergraduate CS program under the liberal arts model:

- An introductory software engineering course introducing the fundamentals
- An introductory Operating Systems Course introducing the fundamentals of OS function, design, and implementation
- An introductory database systems course with an overview of the nature and purposes of database systems and an introduction to data modeling.

At the end of this course, students must deliver a product with a major software component. The product must demonstrate aspects in the design, analysis, implementation, testing, and evaluation. In our approach, we distribute those four aspects over sixteen weeks with the following milestones:

- Definition of the problem and its objectives
- Management of the project and teamwork
- Literature overview
- Alternatives for design and software and or project design methodology
- Specifications: software/project requirements
- Modeling and analysis
- Early-release prototyping

- Mid-semester software/project documentation
- Mid-semester checkpoint presentation
- Implementation
- Testing and verification
- Critical appraise: the analytical process for the systematic evaluation and validation of the software product
- Documentation
- Final product demonstration: Final software/project delivery and oral presentation

The above milestones are carefully scheduled by the project team and overseen by the supervisor over the course of the semester. Close follow-ups, by the team and monitoring by the supervisor, are necessary for the thorough completion of requirements. Indeed, enabling the evaluation of the project aspects, including the proposed milestones, is the aim of the target assessment criteria and the overall framework.

3.2 The assessment criteria

Although CPs are of joint efforts of a team of students, any developed assessment framework must allow for sound evaluations of contributions per student. Assessment of student contribution carefully considers the contents and their quality, timeliness of achievements, and professionalism. A variety of assessment tools can be used for CPs and student evaluations. Moreover, assessment tools are deployed to measure indicators within specific criteria. For an increased reliability of measurements, tools are carefully selected to enable the following multiple sources of measurements:

- Project proposal
- Supervisor consultations
- Mid-semester reports on project progress
- An oral final exam per student
- A report delivered upon project completion
- Project presentation examination by a committee

The project examination committee is mainly formed of professor from within the department that the students belong to. In some instances, external examiners are invited from outsides the department and the university; however, all examiners are university professors in areas related to the project topic. In this paper, we propose the following assessment CPs criteria; some names are inspired by those presented in [2]:

- a) Content 60%)
- b) Impact of the CP on the social good (5%)
- c) Integrity and ethical and legal implications (5%)
- d) Project management and teamwork skills (10%)
- e) Written communication (10%)
- f) Presentation and oral communication (10%)

The developed six criteria, A through F, are carefully selected to cover all intended CP aspects within its indicators and map to all SOs at the program level. The most significant part of CP assessment is under Criterion A that measures the level of project contents achievement. The contents comprise reviewing the literature, the design and implementation techniques, use of technology, analysis and evaluation, and the identification of future work. Furthermore, a weight of 60% of the CP evaluation is allocated for Criterion A (See Table 1). Criterion B assesses the extent to which the proposed CP addresses a challenging problem for the social good. Criterion C ensures that students are clearly aware of the ethical and legal implications surrounding human subject data. Moreover, Criteria C through F captures a bouquet of CP requirements that comprises the understanding of legal implications and adherence to professional ethics, project management aspects, and documentation and presentation skills. Table 1 presents the list of criteria, indicators, and the allocated weights.

 Table 1. The framework hierarchy: criteria, performance indicators, in addition to the allocated weights; Criteria D, E, and F are inspired by the work presented in [2].

A Content (Total %60, Supervisor; %35, Examination Committee %25)
1.Summary, comparisons and evaluations of various concepts, research findings and current theories and
models in core content areas of computer sciences and computing literature in general. (10%)
2.Identification of principles and techniques that are relevant to the project and ability to apply them within specific problem domain. (10%)
3.Interpretability, creativity, and originality of the adopted methodology and the developed solution. (10%)
4. Alternative research solutions, if applicable, and benchmarking with competitors. (10%)
5. Identification, mastering, and use of hardware/software tools. (10%)
6.Robustness and interpretability of results. (5%)
7.Identification of further improvements and future work. (5%)
B. Social Impact (%5, Supervisor and Examination Committee Members)
1.Addressing a problem that stems from a social need, and thus has a social impact. (2%)
2.Adaptability and use by people without an engineering/computing/technology background. (2%)
3.Evaluation of computing solutions that consider global and/or societal factors relevant to the region. (1%)
C Integrity and ethical and legal implications, and regulations (%5, Supervisor and Examination Committee Members)
1.Demonstration of scientific and professional ethics, especially in the interaction with team members and advisor. (2%)
2.Adherence to legal principles, rules, and code of conduct. (2%)
3.Attainment of the criteria and abiding by the regulations that govern the project and/or the data used in the project1. (1%)
D Project Management and Teamwork skills (%10, Supervisor)
1.Ability to work individually, or as part of the team where appropriate, to formulate, analyze, design, and implement a significant project (3%)
2.Contribution to the team project/work. (3%)
• • • • • • • • • • • • • • • • • • •

¹Whenever applicable, adherence to IRB guidelines involving human research subjects.

3.Taking responsibility. (4%)
E Written Communication (%10, Supervisor and Examination Committee Members)
1.Organization and logic. (4%)
2. Writing style (word choice, grammar and sentence structure. (4%)
3.Use of References. (2%)
F Presentation and Oral Communication (%10, Supervisor and Examination Committee Members)
1.Mechanics. (2%)
2.Organization. (2%)
3.Delivery. (2%)
4.Relating to audience. (2%)
5.Response to questions. (2%)

3.3 Bridging capstone projects and outcomes at the program level

The framework's set of carefully developed performance indicators enables a variety of measurements of CP outcomes at the course levels. To benefit from CP measurements in the program review process, the framework is built upon the adoption of ABET 2019 complete set of SOs as both the program and the course outcomes. The adoption of the complete set of SOs as CLOs guarantees the coverage of required outcomes within a CP course. With no doubt, such a unified arrangement of CLOs and SOs facilitates closing the continuous improvement of ABET's Criterion 4. In Table 2, we present the mapping among assessment criteria and ABET SOs. Indeed, the relationships between CP indicators and SOs are many-to-one, where several indicators' scores are aggregated to measure the attainment of a single SO. The following represent our CP CLOs that are the same as the newly developed ABET SOs for Computer Science Programs:

- "Analyze a complex computing problem and to apply principles of computing and other relevant disciplines to identify creative and original solutions.
- Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.
- Communicate effectively in a variety of professional contexts: example presentations and software documentation
- Recognize professional responsibilities and make informed judgments in computing practice based on legal and ethical principles.
- Function effectively as a member or leader of a team engaged in activities appropriate to the program's discipline.
- Apply computer science theory and software development fundamentals to produce computing-based solutions."

SOs	CP Indicator
1	A2, A6
2	A1, A3, A4, A5, A7
3	E1, E2, E3, F1, F2, F3, F4, F5
4	C1, C2, C3
5	D1, D2, D3
6	A2, A3, B1, B2, B3

Table 2. Criteria mapping to ABET SOs.

Adopting ABET SOs as the intended CLOs of the capstone course enables multiple mutual benefits. First, the adoption of SOs as CLOs unifies their assessment without the need for mapping them to each other and accordingly the need to develop an additional statistical aggregation. Second, such an adoption guarantees the literal coverage of all SOs within the course intended outcomes. Accordingly, the framework maintains a 3-level hierarchy of evaluation metrics, namely, criteria, indicators, and their rubrics.

3.4 Hierarchy of evaluation metrics: criteria, indicators, and rubrics

The suggested framework is a three-level hierarchy. The top-most level is the set of criteria that covers all intended CP aspects. A rich set of indicators stems from the criteria to specify the intended measurements. The bottom-level of the framework includes extensive set of analytic rubrics for each indicator. The rubric descriptors specify the quality scale of achievements and the intended requirements. The developed rubrics and the statistical formulations are presented in Section4.

4 Measurement Rubrics

To further develop the proposed framework, analytic scoring rubrics are carefully created based on the specific requirements of the intended CP context. We base our analysis around twenty-Four indicators that map onto the set of revised ABET SOs (See Table 2). Furthermore, the adopted scale of rubrics consists of four attainment levels: a beginning level (B), a developing level (D), a competent level (C), and an accomplished level (A).

To verify the suitability of the created rubrics, we have consulted with four Professor of Computer Science, besides, comparing with rubrics from [2,3,14,15]. The aim of the developed rubrics includes adopting solid descriptors and a variety of performance levels. These four levels in turn correspond to a range of percentages given by [40-69, 70-79, 80-89, 90-100] respectively. The selected ranges of percentages are for a scale that considers 70% as the starting point of the D level. With this level of granularity, we can assess the deliverables of the CP at any of the criterion or indicator levels or the combination of the two. We adopt the following weighted average to aggregate all indicators: $S = \sum_{j=1}^{J} \omega_j p(l_j)$. where S is the combined score, $p(l_i)$ is

the score percentage of the j^{th} indicator I_j and ω_j is the weight of the j^{th} performance indicator I_j such that $0 < \omega_k < 1, j = 1, ..., J$ and $\sum_{j=1}^{J} \omega_j = 1$, where J is the number of performance indicators i.e. in our case, J = 24. The weights are described in Table 1.

Capstone project assessment criteria cover aspects related to global and regional social impacts; understanding of integrity and ethical issues, legal implications, and regulations; management; and effective communication skills. For some part of the created indicators, we are inspired by the rubrics developed in [2, 14, 15]. Although our framework is primarily created for assessing computer science CPs, minor modifications enable its use in similar disciplines.

In Tables 3 through 8 we delve into many details surrounding Criteria A through F. Each row in these tables represents a certain relevant mapping to an ABET outcome. The various columns in each table describe the attainment per scale point (Beginning, Developing, Competent, or Accomplished). We present an overview of the content of these tables as follows. Criterion A in Table 3 carries with it a significant percentage of the overall score, as it addresses several focal points related to the overall quality of the project. The stated rubrics cover the various stages of the project from beginning to end. The indicators dwell on evidence of a thorough literature survey, of a robust understanding of computing principles and techniques, and of a sound methodology that is interpretable and yet creative and original. The indicators investigate whether there has been enough exploration of alternative research solutions, and to what extent the team has conducted benchmarks against competitor solutions.

Criterion A is also concerned with the extent to which appropriate hardware and/or software tools have been exploited, and finally, with the level by which results are interpretable and future work and improvements are identified. Beginner levels are those mostly lacking in all these indicators whereas competent levels meet those criteria and beyond.

Criterion B in Table 4 probes into the social impact that the proposed CP aims to address. The indicators investigate the extent to which the project addresses issues of social impact, examples of which can include poverty, education, or crisis management, to name a few, and the extent to which the team exhibits awareness of the high risks and/or low resourced settings surrounding our society. We then carry on having more indicators of this criterion investigate the level by which the project can be adapted for use by people without the relevant technical background, and the extent to which the team evaluates computing solutions that consider global and regional societal factors. Competent students are expected to demonstrate exemplary awareness of societal challenges and to offer solutions of high impact that are of utility and can be adopted by lay people from all walks of life.

Criterion C in Table 5 assesses the understanding and application of integrity and ethical and legal implications, and regulations. To that end, competent students exhibit consideration and compliance with professionalism and integrity, especially with team members and advisor. Moreover, a team member is to abide by the regulations that govern the project, its data, and show proper adherence to handling guidelines.

Criterion D in Table 6 is aimed at assessing the management skills within the team work as well as the level to which the student has individually contributed to the

project and taken responsibility for sub-tasks. Also, the indicators tackle the time management skills required for achieving major milestones in a timely fashion. The indicators stipulated under Criterion D require the assessment of the project supervisor exclusively, as external examiners have no way to monitor those aspects of the project. Competency at this criterion requires that a student demonstrate active participation in the project as well as a strong initiative leading up to monumental ideas in a timely fashion.

Written communication is addressed in Criterion E in Table 7. The rubrics begin by addressing the organization and the logical order and coherence of ideas. A competent student exhibits a solid logical rationale and a smooth transitioning among ideas as well as a highly relevant body of information. The rubrics then examine the writing style involving the choice of words as well as the grammatical proficiency and the readability of the written document. A competent student has a compelling writing style that captivates the reader throughout. The rubrics finish off with examining the use of references and the level to which the writer provides citations in the text, is accurate in referring to the citations, and chooses relevant and impactful literature references.

Table 8 presents the last criterion assessing the oral and presentation communication skills, as evident from the student's own slides as well as their delivery of the presentation. Particularly, we pay attention to the mechanics manifested in the slides as well as the extent to which they are effectively written, the sequencing and pace of topics in the presentation. We also examine the actual delivery including voice and tone, as well as body language. The engagement with the audience and the level by which the response to questions is appropriate are also address. A student is competent at this criterion if they present extremely creative and well written slides in an engaging manner and show confident presence on stage and excellent engagement with the audience and can navigate through and adapt the presentation considering real time response from the audience.

A weighted aggregation of indicator scores produces the overall percentage grade per student. In addition, simple averaging using the indicator mapping presented in Table 2 enables the calculation of attainment scores per indicator, criteria, and student outcome. In Section 4, the benefits, challenges, results of deployment, and validation of the proposed framework are investigated.

[Mapping to ABET outcomes]	Beginning	Developing	Competent	Accomplished
Summary, compari- sons and evaluations of various concepts, research findings and current theories and models in core content areas of computer sciences and computing literature in general. [SO2]	Literature review is poor in its time-span coverage, quality of publishing venues, as well as breadth and depth of topics. Relevance of topics discussed is poor. Attempt to support assertions with evidence is poor. Content is exces- sively marred with errors.	quality of publishing venues, as well as breadth and depth of topics. Relevance of topics discussed is satisfactory. Attempt to support assertions with evidence is satisfactory. Content	as well as breadth and depth of topics. Relevance of topics discussed is good. Attempt to support assertions with evidence is good.	Literature review is excellent in its time- span coverage, quality of publishing venues, as well as breadth and depth of topics. Relevance of topics discussed is excellent. Attempt to support assertions with evidence is excellent. Content is occasionally marred with errors.
Identification of principles and techniques that are relevant to the project and ability to apply them within specific problem domain [SO1, SO6]	Demonstrates a basic understanding of principles; fails to apply them within specific problem domain.	Includes basic principles and techniques relevant to project but misses some others. Fails to develop complete theoretical or design framework for the project.	Provides good computational and technological framework for the project; applies principles and techniques correctly to problem domain.	Renders that project completely ground- ed in computational principles and technologies; applies them to problem correctly and clearly establishes their relevance.
	The interpretability of the methodology is poor. The creativi- ty of solutions is poor. The novelty of the proposed work is completely lacking and there is no impact observed to the work proposed.	The interpretability of the methodology is satisfactory. The creativity of solu- tions is satisfactory. The proposed work has some novelty and there is some mild impact ob- served to the work proposed.	The interpretability of the methodology is good. The creativ- ity of solutions is good. The proposed work has evident novelty and there is adequate impact observed to the work proposed.	The proposed work has substantial novelty and there is an extremely prom- ising impact ob-
tors.	Presents only one alternative research solution or gives clearly infeasible alternatives. Omits reference to compet- itors entirely.	Experiences short- comings in explor- ing and identifying alternative research solutions. Attempts to refer to competi- tors but omits benchmarking the results.	Identifies alternative solutions to some fair degree and attempts to bench- mark against com- petitor solutions.	Achieves the final design after review- ing reasonable alternatives. Presents comprehensi- ble/conclusive benchmarks.
mastering, and use of hard- ware/software tools	Suffers from serious deficiencies in understanding the correct selection and/or the mastering and use of hardware and software tools.	Demonstrates mini- mal application, mastering, and/or use of appropriate hardware and soft- ware tools.	Masters hardware and software tools and uses them with effectiveness to develop designs. Further improve- ment could be made.	Masters hardware and software tools and uses them highly effectively to devel- op and analyze designs. Final prod- uct is highly profes- sional.

Table 3. Criterion A (Content); rubric is partly inspired by the tool presented in [2].

interpretability of results. [SO1]	experiments and tests are inconclu- sive: results are	results are inconclu- sive but not usable for further investiga- tion. Attempts to	Testing is adequate; analysis and results are acceptable, complete, and	analysis and results are robust, usable and highly interpret- able even to a non-
Identification of further improve- ments and future work. [SO2]	No direction for further improvement is provided.	One or two ideas for future expansion are listed but may not be practical.	which one or two are practical and ade- quate, for further improvements are	Several novel direc- tions for important expansions of the current ideas are thoroughly ex- plained.

 Table 4. Criterion B (Impact of the CP on the Social Good); rubric is partly inspired by the tool presented in [2].

[Mapping to ABET outcomes]	Beginning	Developing	Competent	Accomplished
thus has a social impact.	No addressing of needs on issues of social impact, ex- amples of which can include poverty, education, or crisis management, to name a few. No awareness of the high risks and/or low resourced settings surrounding our society.	education, or crisis management. Lim- ited awareness of the high risks and/or low resourced settings surrounding	issues ranging of social impact, ex- amples of which can include poverty, education, or crisis management. Exhib- its awareness of the high risks and/or low resourced	Addresses in an impactful manner needs on issues of great social rele- vance, examples of which can include poverty, education, or crisis manage- ment. Exhibits conscientious awareness of the high risks and/or low resourced settings surrounding our society.
background. [SO6]	for use by people	engineering, compu-	for use by people without an engineer-	Highly adapts for use by people with- out an engineering, computing, or technology back- ground.
Evaluation of com- puting solutions that consider global and/or societal factors relevant to the region. [SO6]	No evaluation of computing solutions that consider global and/or societal factors relevant to the region.	computing solutions that consider global and/or societal factors relevant to	Evaluates computing solutions that con- sider global and/or societal factors relevant to the region.	Efficiently evaluates computing solutions that consider global and/or societal factors relevant to the region.

[Mapping to ABET outcomes]	Beginning	Developing	Competent	Accomplished
Demonstration of scientific and pro- fessional ethics, especially in the interaction with team members and advisor. [SO4]		understanding but still complies with principles of scien- tific, professional and/or academic integrity, especially	Exhibits understand- ing and compliance with principles of scientific, profes- sional and/or aca- demic integrity, especially with team members and advi- sor.	Clearly documented understanding of compliance with all relevant ethical guidelines; clearly establishes author- ship of the project work.
Adherence to legal principles, rules, and code of conduct. [SO4]	ing of legal princi-	still complies with legal principles and/or implications.		all relevant legal guidelines and implications.
Attainment of the criteria and abiding by the regulations that govern the project and/or the data used in the project. [SO4]	poor abiding by the	tion of abiding by the regulations that govern the project and/or the data used in the project. Whenever applica- ble, some attempt and recall of IBB	Exhibits attainment of the criteria and abiding by the regulations that govern the project and/or the data used in the project. Whenever applica- ble, shows adher- ence to IRB guide- lines governing human research subject.	Strict and explicit reference in the project towards the attainment of the criteria and abiding by the regulations that govern the project and/or the data used in the project. Whenever applicable, diligently assures the reader of the strict adherence to IRB guidelines governing human research subjects.

Table 5. Criterion C (Integrity and Ethical and Legal Implications); rubric is partly inspired by the tool presented in [2].

[Mapping to ABET outcomes]	Beginning	Developing	Competent	Accomplished
Ability to Work indi- vidually, or as part of the team where appro- priate; to formulate, analyze, design, and implement a signifi- cant project. [SO5]	ed standards. By and large always falling behind	operating within the team. Contributions to the project are not significant despite exceeding minimal requirements.	satisfactory. Can follow a timeline yet falls behind when taken by other commitments. Deliverables do not contain any faults but leave room for	Is instrumental in leading the team. Contributions to the project are signifi- cant. Strictly abides by a timeline. Deliv- erables are well- formulated, de- signed, and imple- mented
Contribution to the team project/work [SO5]	Individual contri- butions to the team are not relevant or useful, and do not address the team's needs;	Only when prompt- ed, embarks on contributing infor- mation to the team. Tries to provide some ideas but suggestions are not sufficiently devel- oped to meet the team's needs.	Can provide some basic and useful information to assist in the project and occasionally makes some useful sugges- tions to the team that meet its needs.	tematically offers well developed and clearly expressed
Taking responsibility [SO5]	Takes no respon- sibility whatsoev- er and shows no initb ? 'iative at all relyin /g on the other team members to do the work. By and large misses meetings and when present demonstrates marginal partici- pation.	Can perform as- signed tasks but regularly needs reminders and prompts. Delegates the challenging parts of the project to others. Does not have a constructive presence during meetings.		Is able to perform all assigned tasks highly effectively. Takes initiative in setting up meetings and is the lead participant in the discussions that take place then.

 Table 6. Criterion D (Project Management and Teamwork Skills); rubric is partly inspired by the tool presented in [2].

[Mapping to ABET out- comes]	Beginning	Developing	Competent	Accomplished
Organization and logic [SO3]	The information in the text has no logical order, lacking in many important details, and is difficult to under- stand.	The text exhibits a weak organization. The presentation of ideas lacks coherence and shows no smooth transition between them.	The text exhibits a reasonable organiza- tion. The information in the text has some- what a logical back- ing and an attempt to provide a project rationale is made.	The written report exhibits strong clarity and a solid logical rationale. Transition- ing among ideas is smooth. The infor- mation presented is at large very relevant and thorough, all resulting in a highly informative piece of text.
Writing style (word choice, grammar and sentence structure) [SO3]	Choices of words and expressions often misleading. Text suffers from numerous errors in grammar that compromise on the clarity of the docu- ment.	Choice of words and level of grammatical proficiency is gener- ally adequate. Yet, the document still is difficult to read.	The writing style and the general flow of the text are satisfacto- ry.	The writing style is compelling and is able to captivate the reader till the end.
Use of Refer- ences [SO3]	Most references in- cluded are inaccurate and are not relevant. Almost inexistent attempt to provide citations in the text.	Most references provided are clearly indicated but have little impact in the literature. A con- servative attempt to provide citations in the text is made.	Prior work is proper- ly cited in most places where needed (e.g. when referring to theories, assump- tions, and findings). Minor exceptions exist. References are accurate in referring to author names, journals or proceed- ings, volume num- bers, page numbers, and year of publica- tion. References have a modest impact in the literature.	Prior work is properly cited in most places where needed (e.g. when referring to theories, assumptions, and findings), with no exceptions whatsoev- er. References are accurate in referring to author names, journals or proceedings, vol- ume numbers, page numbers, page numbers, and year of publication. Refer- ences have a great impact in the litera- ture.

Table 7. Criterion E (Written Communication); rubric is partly inspired by the tool presented in [2].

[Mapping to ABET outcomes]	Beginning	Developing	Competent	Accomplished
Mechanics [SO3]	Content of slides seems to be completely irrele- vant, reflecting a lack of understanding how the presentation should be crafted. Numerous mistakes appear in the presentation's text. Speaker largely unsure how to flow from one slide to the next.	Slides are boring and largely ineffective, despite being largely error free.	Slides are generally good and convey key messages reasonably well.	Slides are extremely creative and expose the main aspects behind the projects. Audience remains interested throughout the presentation.
Organization [SO3]	Sequencing and pace of topics in the presenta- tion seems to be in a complete disarray, making it difficult to derive any clear conclu- sions.	Presentation oscillates between sometimes following an organized track and some other times not so. In gen- eral, though, one is still able to derive plausible conclusions.	along; the key points are present- ed; leads up to	The presentation is clear and slowly builds up in a fo- cused manner. De- tails in the presenta- tion are entirely relevant and help the audience derive a deep understanding of the topic.
Delivery [SO3]	Speaks in an extremely low voice and mumbles. Too many filler words that distract the audi- ence	Speaks in a low voice but only occasionally inaudible. Uses some distracting filler words but mostly articulates in a modest manner.	Speaks in a clear voice and delivers a generally effective presentation.	Speaks naturally and in a confident man- ner. Goes beyond merely conveying the message but over to enhancing it with the help of an effective tone, pitch, and body language.
Relating to audience [SO3]	Reads most of the presentation from the slides or note. Fails to maintain eye contact or maintain a catching body language. Com- pletely oblivious to audience reactions.	Shy attempts to main- tain eye contact and to move around in a catching manner. A modest attempt to improvise aside from the notes written in the presentation. Somehow aware of the presence of audience (at the very least, those sitting closely). Very brief and somehow dis- missive response to audience questions.	around in a catching manner. Is able to deliver the presenta- tion as if conversing with the audience.	Maintains excellent engagement with the audience throughout the audience. Is able to modify the presenta- tion on the spot as needed based on audience engagement, questions, and com- ments.

Table 8. Criterion F (Presentation and Oral Communication); rubric is partly inspired by the tool presented in [2].

Response to questions [SO3]	Generally unprepared to answer questions. Misunderstands most of the questions and fails to provide appropriate response.	answer questions. Understands most of the questions but demonstrates difficulty in providing correct or well-informed re-	Demonstrates a clear understanding of the questions and is well prepared to answer them. Pro- vides mostly correct and well-informed responses.	and respond with ample information beyond what's needed.
-----------------------------------	--	--	--	--

5 Analysis and Evaluation

This section presents the result and evaluation of the proposed framework. In addition, the benefits of the proposed assessment framework and its challenges are discussed. The framework uses an assessment structure, formulation and scoring like the one used in [16]. In addition, several CPs covering an extensive selection of computer science problems are used to evaluate the assessment tool. The tool is also evaluated with several evaluators who have different experience and level of expertise in different areas.

5.1 Case-study setup

The presented study includes seven projects with a total of 25 students and a typical project team size of 3-4 students. Each project has a single supervisor and is examined by a committee of four professors. The examination committee evaluates the project report, presentation, and the developed prototype during a demonstration. The supervisor examines all deliverables. The target passing score per student is 70% for the overall course grade; the same percentage is adopted as the target score of SOs.

A pilot study for a single program is applied to evaluate the proposed framework. The evaluated CP's are from an institution of higher education which adopts the American model of higher education. Data collection started with refining the documents and forms already in use at the institution. This led to balanced artifacts. The analytic version of the tool was then used to collect data. Training sessions were then conducted for the evaluators to ensure evaluators understood the rubrics. Then data collection was conducted after calibration.

5.2 Analysis

As shown in Tables 9 and 10, most of the indicators and SOs were not different before and after calibration. Even with the indicators which had different scores before and after calibration the 4-point rubric scale mapping did not change. This indicates that the Rubric is clear enough to be used by evaluators before and after calibration without the fear of changing the 4-point score of any of the indicators. Based on the

rater calibration, several improvements are identified. The improvements include the following:

- Add the identifier "ACM" for core areas in indictor A1 to specify the target core areas. ACM core areas are detailed in [17].
- Add "during the design phase" to indicator A2 to identify the specific stage of development.
- Remove "impact" from indicator A3 to avoid overlapping measurements with indicator B1.
- Specify the code of conduct in C2 as the "Code of conduct of the Institution"
- Modify "good" under F1 under Competent by "well-designed" to better match the intended meaning by the rubric designers.

The measurements made for the proposed indicators identified additional opportunities for improvement on the SO level (See Table 10), namely SO4 and SO6. The identified improvement is on the student abilities to recognize professional responsibilities and make informed judgments in computing practice based on legal and ethical principles. In addition, the needed improvement is related to the student effective application of computer science theory and software development fundamentals to produce computing-based solutions. At the program level, the attainment scores are usually combined with triangulated measurements from other courses to reach a final attainment score and improvement decisions.

As presented in Section 3, the proposed framework enables evaluating student performance as the weighted sum of indicator scores. The results of the evaluated 25 students ranged between 70% and 90%. Although all students met the projected passing grade of the course with their overall score, the tool allowed for the identification of improvements to their intended abilities at the indicator, criteria, and SO levels.

Upon incorporating the suggested modifications to the rubric, the proposed framework is ready for deployment within Computer Science Programs that are aligned with ABET requirements and ACM recommendations. However, tuning and customizing the framework is straightforward. Customizations can be applied to the criteria, indicators, rubrics, and the choice of the aggregating statistical formulation.

5.3 Evaluation

Several benefits exist for the framework proposed in this paper. The framework limits bias and uncertainty of evaluator's thereby promoting quality in assessment. This is due to the clear measurement structure resulted from the conceptual basis of the framework. At the program level, a main source of measurement are the integrated key CP indicators. In addition, conclusions at different levels of abstraction are possible using the framework. Measurements are made at the indicators, criterion as well as ABET SO levels of abstraction. The measurement structure and statistics in the framework are not limited to computer science and can be applied for any other discipline. Indicators and rubrics are comprehensive and rich yet simple enough to be understood by faculty members as well as students. A smooth transition in descriptors

is used in the rubrics which makes it easier for evaluators to pick the appropriate descriptor.

However, the framework poses several challenges to its implementation. A prerequisite for implementing the framework is that a culture of assessment be present. The evaluators should be first trained on how to use the rubrics. They should also be committed to reviewing thoroughly all artifacts being measured such as presentations, reports and essays. In addition, the time constraints placed on the evaluation process are also considered as a great challenge. To that end, the allocation of four examiners per project can be reconsidered and replaced by a smaller evaluation committee.

6 Conclusion

Computer Science programs rely heavily on senior CPs to demonstrate student abilities accumulated throughout the program. CPs are rich in requirements and deliverables; this makes them of unique importance in evaluating student performance and assessing their attainment of SOs. In this investigation, we present a framework for systematically, accurately, effectively, and jointly evaluating student performance, assessing learning outcomes, and accordingly assessing SOs. The hierarchy of the developed framework is of three levels that comprise criteria, indicators, and an extensive set of analytic rubrics. A single-institution pilot study is executed to calibrate the proposed rubrics. The study includes several CPs from wide spectrum of computer science topics. The tool tuning attained a small variance in scores after the calibration of raters. A difference of 2% is found between the scores—before and after calibration. The proposed framework is easy to deploy and was found to effectively eliminate subjectivity in assessment and evaluation. Future work includes carrying out a study that involves multiple programs.

Indicators	Before Calibration (BC)	After Calibration (AC)
A.1 (10%)	32 (B)	32 (B)
A.2 (10%)	100 (A)	76 (A)
A.3 (10%)	72 (D)	72 (D)
A.4 (10%)	52 (B)	56 (B)
A.5 (10%)	100 (A)	100 (A)
A.6 (5%)	84 (C)	84 (C)
A.7 (5%)	60 (D)	60 (D)
B.1 (2%)	44 (B)	44 (B)
B.2 (2%)	60 (D)	60 (D)
B.3 (1%)	16 (B)	16 (B)
C.1 (2%)	60 (D)	60 (D)
C.2 (2%)	16 (B)	16 (B)
C.3 (1%)	72 (D)	72 (D)
D.1 (3%)	84 (C)	84 (C)
D.2 (3%)	84 (C)	84 (C)

 Table 9. Assessment results in percent and their corresponding scale point: Beginning (B), Developing (D), Competent (C) and Accomplished (A).

D.3 (4%)	84 (C)	84 (C)	
E.1 (4%)	56 (B)	56 (B)	
E.2 (4%)	60 (D)	60 (D)	
E.3 (2%)	32 (B)	32 (B)	
F.1 (2%)	84 (C)	84 (C)	
F.2 (2%)	100 (A)	100 (A)	
F.3 (2%)	100 (A)	100 (A)	
F.4 (2%)	100 (A)	100 (A)	
F.5 (2%)	84 (C)	84 (C)	
Weighted Average	70.32%	68.32%	

 Table 10.
 Mapping scores to ABET SOs before and after calibration.

SOs	CP Indicator	BC	AC
1	A2, A6	92%	80%
2	A1, A3, A4, A5, A7	83.2%	78.4%
3	E1, E2, E3, F1, F2, F3, F4, F5	77%	77%
4	C1, C2, C3	49.3%	49.3%
5	D1, D2, D3	84%	84%
6	A2, A3, B1, B2, B3	58.4%	53.6%

7 References

- Bachnak, R., An approach for successful Capstone projects. Proceedings Frontiers in Education 35th Annual Conference, Indianopolis, IN, 2005, pp. F4D-18
- [2] Yousafzai, J., Damaj I., El Abd M., A unified approach for assessing capstone design projects and student outcomes in computer engineering programs, 2015 IEEE Global Engineering Education Conference (EDUCON), Tallinn, 2015, pp. 339.<u>https://doi.org/10.11</u> 09/EDUCON.2015.7095993
- [3] Damaj, I., Yousafzai J. (2019). Effective Assessment of Student Outcomes in Computer Engineering Programs using a Minimalistic Framework. International Journal of Engineering Education, Tempus Publications, 35(1.A): 59-75
- [4] Olarte, J. J., Domi'nguez, C., Jaime, A., F. Garcı'a-Izquierdo J. (2016), Student and Staff Perceptions of Key Aspects of Computer Science Engineering Capstone Projects. IEEE Transactions on Education, 59(1): 45-51. <u>https://doi.org/10.1109/TE.2015.2427</u> 118
- [5] Pérez, C. D., Elizondo, A. J., García- Izquierdo, F. J., Olarte Larrea, J. J. (2012). Supervision typology in computer science engineering capstone projects. Journal of Engineering Education, 101(4): 679-697. <u>https://doi.org/10.1002/j.2168-9830.2012.tb01124.x</u>
- [6] Ahmad, E., Raza, B., Feldt, R., Assessment and Support for Software Capstone Projects at the Undergraduate Level: A Survey and Rubrics. Frontiers of Information Technology, 19-21 December 2011, Islamabad, Pakistan. pp. 25-32. <u>https://doi.org/10.1109/FIT.2011.13</u>
- [7] Bihari, T., Malkiman I., Chaabouni M., Bolinger J., Ramanthan J., Ramath R., Herold M. Enabling Scalability, Richer Experiences and ABET-Accreditable Learning Outcomes in Computer Science Capstone Courses through Inversion of Control. 41st ASEE/IEEE

Frontiers in Education Conference, October 12-15, 2011, Rapid City, SD. https://doi.org/10.1109/FIE.2011.6142872

- [8] Chen, C., Hong, Y., Chen, P. (2014). Effects of the Meetings-Flow Approach on Quality Teamwork in the Training of Software Capstone Projects. IEEE Transactions on Education, 57(3): 201-208. <u>https://doi.org/10.1109/TE.2014.2305918</u>
- [9] Pérez, C. D., Elizondo, A. J., García-Izquierdo, F. J., Olarte Larrea, J. J. (2012). Supervision Typology in Computer Science Engineering Capstone Projects. Journal of Engineering Education, 101(4): 679–697. <u>https://doi.org/10.1002/j.2168-9830.2012.tb01124.x</u>
- [10] Utesch, M. C., Hauer, A., Heininger, R., & Krcmar, H. (2017). Automated Stock Trading-Developing the Serious Game FSTG to Teach the Topic of Finite State Machines. International Journal of Engineering Pedagogy, 7(1). <u>https://doi.org/10.3991/</u> <u>ijep.v7i1.6524</u>
- [11] Campbell, B., Voelker, J., & Kremer, C. (2015). An analysis of engineering educational standards and outcomes achieved by a robotics summer camp experience. International Journal of Engineering Pedagogy, 5(4), 12-21. <u>https://doi.org/10.3991/ijep.v5i4.</u> 4713
- [12] Baumgartner, I. (2014). A Set of Best Practices to Design Face-to-face Teaching Sessions for Technology-centered University-level Computing Courses. International Journal of Engineering Pedagogy (iJEP), 4(4), 59-66. <u>https://doi.org/10.3991/ijep.v4i4.4000</u>
- [13] ABET CAC Criteria 2018 2019, <u>http://www.abet.org/accreditation/accreditationcriteria/cac-18-19/, Accessed April 4, 2019</u>
- [14] CSE Program: Objectives, Outcomes, Assessments, Program Improvements, The Ohio State University, Accessed October 10, 2019. <u>https://cse.osu.edu/content/cse-program-objectives-outcomes-assessments-program-improvements</u>
- [15] Learning Outcomes Assessment, University of Idaho, Accessed October 10, 2019. <u>http://www.webpages.uidaho.edu/ira/assess</u>
- [16] Damaj, I., Ater Kranov A. (2017). Sustainable Practices in Technical Education: A Quality Assurance Framework. International Journal of Engineering Education, 33: 1627-1642
- [17] ACM Curricula Recommendations, Curriculum Guidelines for Undergraduate Programs in Computer Science, Accessed October 10, 2019. <u>https://www.acm.org/binaries/ content/assets/education/cs2013 web final.pdf</u>

8 Authors

Fatima K. Abu Salem received her BS and MS degrees in Mathematics from the American University of Beirut, Beirut, Lebanon, and the D.Phil. degree in Computing from the University of Oxford, Oxford, U.K. She is currently an Associate Professor with the Computer Science Department, American University of Beirut. Her current research interests include computer algebra, parallel computing, and data science for the public good. Email: <u>fa21@aub.edu.lb</u>

Issam W. Damaj, received his PhD in Computer Science from London South Bank University, London, UK. His ME in Computer and Communications Engineering from the American University of Beirut, and his BE in Computer Engineering from Beirut Arab University (BAU), Beirut, Lebanon. He is an Associate Professor with the Electrical and Computer Engineering Department, BAU, where he is also the Director of the Center for Quality Assurance. His research interests include hardware design, smart cities, and engineering education. Email: <u>i.damaj@bau.edu.lb</u>

Rached N. Zantout received his BE from the American University of Beirut, Lebanon in 1988, his MSc from the University of Florida in 1990 and PhD from the Ohio State University in 1994, all degrees being in Electrical Engineering. He is a Professor at the Electrical and Computer Engineering Department of the College of Engineering at Rafik Hariri University, Mechref, Lebanon. His research interests include Robotics, Artificial Intelligence, and Natural Language Processing. Email: <u>zan-</u>toutrn@rhu.edu.lb

Lama A. Hamandi received her BE from The American University of Beirut, Lebanon and her MSc and PhD degrees from the Ohio State University; all degrees being in Electrical Engineering. She is a Senior Lecturer at the Electrical and Computer Engineering Department of the Maroun Semaan Faculty of Engineering and Architecture at the American University of Beirut, Beirut, Lebanon. Her research interests include Parallel Processing, Natural Language Processing, and Computational Linguistics. Email: <u>lh13@aub.edu.lb</u>

Article submitted 2019-10-12. Resubmitted 2019-11-26. Final acceptance 2019-11-27. Final version published as submitted by the authors.