International Journal of Engineering Education Vol. 37, No. 5, pp. 1429–1440, 2021
Printed in Great Britain
0949-149X/91 $3.00+0.00

Applying the Computing Professional Skills Assessment Method to ESL Students in a Computing Program*

MAURICE DANAHER and ANTHONY RHODES
Zayed University, Abu Dhabi, U.A.E. E-mail: maurice.danaher@zu.ac.ae, tonyr1848@gmail.com

ASHLEY ATER KRANOV
Khalifa University, Abu Dhabi, U.A.E. E-mail: kranova@live.com

Professional skills, also known as non-technical, transferable or 21st century skills are recognized worldwide as very important for university graduates. These skills include communication, problem-solving and the ability to function successfully on a multi-disciplinary team. While academic programs worldwide strive to develop a solid professional skills base in students, these skills are notoriously challenging to teach and assess. This paper presents the Computing Professional Skills Assessment (CPSA), a performance assessment of the professional skills learning outcomes identified by the Accreditation Board for Engineering and Technology (ABET), for the computing discipline. The assessment consists of a performance task, a rubric and an implementation method. The CPSA has been developed in an iterative manner and results suggest that it can accurately and consistently elicit and measure the targeted professional skills in computing students within appropriate contexts. Considering two components of the Communities of Inquiry model, cognitive and social presence, specific reliability and validity protocols were developed for the CPSA. These involved a pilot study, iterative development and assessment of the rubric, establishment of interrater reliability, the production of computing specific scenarios which are suitable for non-native English language speakers, and an implementation strategy using asynchronous discussion boards. Based on our findings through more than 50 CPSA implementations over a seven-year period, we have found that the instrument can be used reliably and validly as a measure of the professional skills learning outcomes for English as a second language learners in a computing program.

Keywords: asynchronous discussion board; learning outcomes; professional skills; rubric

1. Introduction

Now globally recognized as important within computing, the professional skills, also known as non-technical, transferable, or 21st century skills, relate to communication, problem-solving, teamwork, ethics and social responsibility, computing's impact on society, and lifelong learning. They have been identified as critical learning outcomes for success in the global work environment [1–3], and since knowledge economies require employees with these skills, this has translated into a demand on educational institutions to produce graduates with these capabilities. However, technical programs in particular have been criticised for being out of date, and not adequately preparing students for the current and future world of employment [4]. Because of this, academic programs have responded by showing an increased focus on professional skills, with the integration of communication, problem-solving, teamwork, ethics and social responsibility, computing's impact on society, and lifelong learning into the curriculum. The problem, however, is that the professional skills have long been seen, and remain, challenging to teach and assess [5, 6].

When it comes to existing assessment instruments, they often only evaluate some of the professional skills or measure them indirectly through methods like interviews, opinion surveys, or focus groups [19]. Effective data-driven decision-making to improve curricula should not rely on methods based on perceptions of learning or that are unrelated to one another. It is essential that direct measures of learning are used so that meaningful data that demonstrate how students actually perform is gathered. It is also preferable that these direct measures are integrated and cover all of the professional skills, so that the assessment process is not cumbersome.

In response to the deficiencies in current tools for assessing the professional skills in computing programs, the Computing Professional Skills Assessment (CPSA) has been developed. Built around the framework provided by the Accreditation Board for Engineering and Technology's (ABET) professional skills, the CPSA is a unique assessment instrument for computing programs which measures student performance outcomes in the professional skills through a method which addresses all of the skills. The components of the CPSA are: (1) a written scenario with instructions and a set of

* Accepted 1 June 2021.

discussion guiding questions; (2) a scenario creation checklist; (3) the CPSA rubric; and (4) the method of implementation itself.

The primary goal of this research project has been to develop a rigorous and robust method of assessing all six ABET professional skills for computing within a single assessment instrument that would be suitable for English as a second language (ESL) computing students. These are students who are not native-speakers of English, but study computing in an English- medium institution. Because reliability and validity are always related to the context, the specific research questions are (1) To what degree are students attaining the expected level of proficiency in the professional skills? (2) In what ways have scores of the CPSA proven to be a reliable measure of professional skills for ESL computing students? (3) In what ways have scores of the CPSA proven to be a valid measure of professional skills for ESL computing students?

2. Literature Review

Professional skills are challenging to assess [5, 6]. However, programs, courses, and student achievement must be measured and assessed for the existence of these skills because both learning and accreditation demands it. Both technical and non-technical programs face difficulties in measuring professional skills [1, 3, 8]. One reason is that each of the professional skills, like other learning outcomes, require conceptual elaboration for the practical purpose of assessment design and implementation [9]. This can be more challenging for the professional skills because they are not as concrete as some technical skills. Another reason is that many technical faculty and program leaders lack experience in assessing professional skills; they may not have been assessed in these skills in their own education or may not have received training in how to teach or measure non-technical skills [10]. Additionally, there are few published assessments that measure computing students' professional skills, especially assessments that assess the professional skills simultaneously.

Though there are limited published assessments measuring computing students' professional skills, assessment methods have been applied to assess delivery and demonstration of the professional skills. Many published measurement methods evaluate the professional skills separately, and as distinct from each other, whereas the professional skills are inter-relational and thus best assessed with tools that can measure more than one at a time. Some of these assessment methods are further problematic as they rely on perceptions, which is an indirect measurement, and so inadequate for valid and reliable assessment of attainment of student learning outcomes. A further constraint of these assessments, in particular portfolios and internships, is that the assessment process may be resource intensive and cumbersome to implement. Poorly performing assessment tools can be harmful to programs because the inadequate or inaccurate data they provide can ill-inform course and program-level decisions [11].

The experience of student internships has been utilized with some success to assess professional skills [12], mainly because all the skills are on view in the work environment. With internships, students are given the opportunity to utilize the full range of professional skills, however assessment can lack control or standardisation if performance is evaluated by the employer. More traditional methods have also been employed, such as take-home written exams [13], which can test the theoretical aspects of whether students understand the use of the professional skills, but not assess the practical application of them dependably. Written exams allow the students to reflect, for example, on a performance task where they may have applied the professional skills, but this does not directly measure or demonstrate their proficiency in this skill. A further form of assessment applied to measure student production of the professional skills has been portfolios [14, 15], but these may rely on reflection, and so are also an indirect measure of the skills.

A method to assess all of the ABET professional skills in engineering was developed at Washington State University's College of Engineering [16]. Known as the Engineering Professional Skills Assessment (EPSA), it measures the professional skill learning outcomes defined by ABET at both program and course level, and is the first method to assess the six ABET professional skills learning outcomes simultaneously and directly. It has inspired the current research project to create the CPSA and has also seeded its development.

The EPSA is a discussion-based performance task which elicits students' knowledge and application of the professional skills. Performance assessment has been shown to be an effective measurement of skills and knowledge exhibited by student participation in a process of creating a final product [17, 18]. Performance assessments are designed to address inter-related learning outcomes and encourage student interaction with authentic issues. The EPSA adopts the typical components of performance assessments, namely: (1) a task to elicit the performance; (2) the performance itself; and (3) a criterion-referenced instrument, such as a rubric, to measure the quality of the performance

[19]. The EPSA consists of a scenario, a student face-to-face discussion of the scenario according to scaffolded prompts, and a rubric for assessing student responses. The scenario briefly describes an authentic and complex engineering issue, similar to a real-life issue faced by engineers in the work place. Students participate in a recorded discussion about the scenario in groups numbering approximately five for 45 minutes. The recordings are transcribed and then analysed by faculty raters using the EPSA rubric. Due to the design of the scenario and the prompts, the students are given the opportunity to demonstrate their abilities in all six skills at the same time. The format of this performance-based task produces actual examples of student skills, thereby increasing the efficacy of the assessment [20].

One of the fundamental differences between the EPSA and CPSA is the move to an asynchronous discussion board. Asynchronous discussion boards have become a cornerstone of many online and blended learning courses [21] because of benefits they can bring. In a meta-analysis into 15 years of discussion boards in higher education, Zhou [22] found that most students talk more in an online discussion than in a face-to-face classroom. This is useful because it demonstrates that the shy, quiet, or more passive learner does indeed engage more online. The elimination of spontaneous response and additional reflective time appears very important. Anderson, Archer, and Garrison [23] also found that this reflection helps to foster critical thinking, an essential skill for a complex world with ill-defined problems awaiting in the work environment. Finally, Salmon [24] showed that co-constructing knowledge and working effectively as a group, which are important aspects of learning, develop quite strongly in online discussions.

3. Theoretical Framework

Whether conscious or not, all teaching and assessment is framed within a theory of learning because the way in which we teach and assess speaks to our understanding of learning. Modern theories of learning have clarified that "assessment practices need to move beyond a focus on component skills and discrete bits of knowledge to encompass the more complex aspects of student achievement" [25, p. 3]. Through cognitive theory we have come to recognize that with assessment we need to know more than what someone knows, we need know how, when, and whether they can utilize the information they know. This requires assessments like the CPSA that are far more complex than traditional tests where the focus is on correct or incorrect responses. We want students to use long-term memory to solve current problems rather than use short-term memory to regurgitate discrete facts that are easily forgotten. More recently in assessment, there has been a growing emphasis "given to social dimensions of learning, including social and participatory practices that support knowing and understanding" [25, p. 3]. Because of this, the Communities of Inquiry model (COI), with its focus on cognitive and social presence has also served as the theoretical underpinning for the CPSA.

The COI model emerged out of research conducted from 1997 to 2001 into computer mediated text-based conferencing, better known as asynchronous online discussion boards today. Research into COI continues to this day and has led to a theory, methodology and instruments [26]. Through a rigorous coding exercise of higher education course discussion boards, Anderson, Archer, and Garrison [23] ascertained the components and structure of the COI. The model is comprised of three core elements, cognitive presence, social presence, and teaching presence, which together form the educational experience. Cognitive presence is the key to success in higher education. It is socially constructed knowledge that is created through sustained communication - for example, the ability to draw conclusions from the discussion. It is nearly impossible to achieve cognitive presence without having established the next core element, social presence. Social presence is the ability of participants to open up and communicate honestly and openly with one another. The final element is teaching presence. This element is normally acted upon by the teacher, but in higher education it is not uncommon for this role to transfer to students. This element is purely supportive in nature. Its purpose is to promote both cognitive and social presence to help attain the educational goals. Since the CPSA is an assessment instrument, teaching presence is extremely limited because it might influence the reliability and validity of the instrument. The purpose of the assessment is to evaluate the performance of students, not the teacher. It is the interplay between cognitive and social presence which impacts upon the development of an effective discussion board and assessment. An effective discussion board assessment will be one in which students are able to demonstrate the learning outcomes, in the case of the CPSA the professional skills, through the demonstration of a high degree of cognitive presence built around social presence. This is important because we know that "much of what humans learn is acquired through discourse and interaction with others". [25, p. 5].

4. Components of the CPSA

The CPSA is a team-based assessment instrument using an asynchronous online discussion forum to have students discuss, analyze, and suggest ways to address complex, multi-faceted, computing-related issues. It is a criterion-referenced assessment which means it is "a type of assessment designed to provide a measure of performance that is interpretable in terms of a clearly defined and delimited domain of learning" [27, p.42]. The components of the CPSA include:

1. Written scenarios with accompanying instructions and discussion questions;
   (a) a scenario creation checklist to aid scenario creation;
2. The scoring rubric;
   (a) the method of implementation itself;
3. An instructors' manual that includes exemplars from discussion transcripts on how to score student work according to the rubric descriptors.

4.1 Scenarios

A scenario is a short article that serves to set the stage for the discussion and provide background information on the topic being discussed. It presents a real-life, multi-faceted, cross-disciplinary computing issue that has both local and global relevance and may contain some technical information. Though scenarios contain trustworthy references, they are not an all-encompassing thesis on the topic; they are meaningful and engaging discussion starters. To be valid, the scenarios need to be written at a level of English that is comprehensible to the students. Otherwise, the task becomes primarily a reading assessment and that is not one of the professional skills. Scenarios have been developed around topics such as cryptography, cyber-attacks, big data, illegal downloading, and information privacy.

Each scenario is preceded by instructions and a set of guiding questions. The instructions provide the context through which the students should proceed with the discussion. The five guiding questions are aligned with the targeted professional skills and designed to frame the discussion and analysis and to facilitate problem solving. The guiding questions not only elicit responses related to each of the professional skills, but they are also the typical questions/steps one should consider when solving a complex problem. Over time, we have learned to stress that the scenarios must be kept short, that the language must be simple, that the instructions are concise, and that the guiding questions are well-aligned to the professional skills. The instructions and guiding questions are as follows:

Introduction
Assume that you have been appointed to a team of 5 or 6 computing professionals within your organization. You have been asked to examine the current issue outlined in the article below. Your team has not been asked to make specific recommendations to solve the problem. Rather, you have been asked to make recommendations that will help the Government decide what next steps they should take.

Prompts
1. What is/are the problem/problems here? Is there an underlying fundamental problem?
2. Who are the major stakeholders and what are their perspectives?
3. What are the major ethical, legal, and security aspects associated with the problem?
4. What are the intended and unintended consequences of existing computing solutions? Consider the consequences on individuals, organizations and society within local and global contexts.
5. What recommendations do you propose that may lead to potential solutions?

4.2 Scenario Creation Checklist

To guide development of scenarios and to ensure they meet the requirements that include appropriate language level, a scenario creation checklist that contains the following criteria is utilized:

- Not focused or dependent on one discipline;
- Complexity;
- Real and relevant problem;
- Context;
- Technical Complexity;
- Elicits engagement;
- References;
- Appropriate for course use (language level).

4.3 The Rubric

The CPSA rubric assesses a sub-section of ABET's student outcomes, the professional skills. Though well-established, the professional skills are quite broad, so a slightly modified version that is a better fit with our assessment task has been created and align to ABET's outcomes (see Table 1.).

In the CPSA rubric, each of the six professional skills is presented on a separate page and includes criteria, descriptors and standardized levels of performance. In order to accurately assess the professional skills, the constructs they represent have been clearly defined. As an example, the rubric component of CPSA 1 is shown in Table 2 (the full rubric

Table 1. Alignment of CPSA professional skills and ABET outcomes=

CPSA Professional Skills Outcomes	ABET Computing Outcomes (2019)
• CPSA 1. Students problem solve from a computing perspective.	b. An ability to analyse a problem, and identify and define the computing requirements appropriate to its solution.
• CPSA 2. Students work together as a group.	d. An ability to function effectively on teams to accomplish a common goal.
• CPSA 3. Students consider ethical, legal, and security aspects.	e. An understanding of professional, ethical, legal, security and social issues and responsibilities.
• CPSA 4. Students communicate professionally in writing.	f. An ability to communicate effectively with a range of audiences.
• CPSA 5. Students analyze the consequences of existing computing solutions within local and global contexts.	g. An ability to analyse the local and global impact of computing on individuals, organisations and society.
• CPSA 6. Students interpret, represent, and seek information.	h. Recognition of the need for and an ability to engage in continuing professional development.

Table 2. CPSA rubric example

CPSA 1. Students problem solve from a computing perspective.

Rater Composite Score for Skill________

Definition: Students define and differentiate between the problems raised in the scenario with reasonable accuracy. Students recommend potential non-technical and technical solutions from a computing perspective. Students identify relevant stakeholders and explain their perspectives.

	0 - Missing	1 - Emerging	2 - Developing	3 - Practicing	4 - Maturing	5 - Mastering
Problem Identification	Students do not identify the problems in the scenario.	Students begin to define the problems. Attempts to define the problems may be general, narrow, and/or inaccurate.		Students define the problems with reasonable accuracy and differentiate between them with limited justification.		Students convincingly and accurately define the problems and differentiate between them, providing realistic justification.
Recommendations for Solutions	Students do not make any recomm-endations for potential solutions.	Students may recommend potential solutions that don't fit the identified problems. Students may make recommendations for potential solutions without identifying the problems first.		Students propose reasonably viable recommendations for non-technical and technical potential solutions.		Students propose detailed and viable recommendations for non-technical and technical potential solutions.
Stakeholder Perspectives	Students do not identify stakeholders.	Students begin to identify stakeholders and their perspectives.		Students explain the perspectives of major relevant stakeholders and convey these with reasonable accuracy.		Students thoughtfully consider perspectives of diverse relevant stakeholders and articulate these with clarity and accuracy.
Comments						

and other components of the CPSA are available at www.cpsa.ae).

Within the rubric, each of the professional skills outcomes is scored on an integer scale from 0 to 5, labelled as *Missing (0), Emerging (1), Developing (2), Practicing (3), Maturing (4) and Mastering (5)*, while descriptors within each outcome align to score levels. The rubric was designed so that the target for students' year of study would align to the labelled integer scale. For example, the target level of performance for first year students is *Emerging (1)*, and this increases incrementally until master's students are expected to achieve *Mastering (5)*. In a recent study that looked at this specifically for 3rd year, 4th year, and master's cohorts this desired alignment was found to exist [28].

4.4 The Method of Implementation

In terms of the CPSA method, since it is not a typical exam, project, or lab assignment, most students have no or little experience with this type of assessment. Because of this, we implemented a two-stage approach. The first stage of implementation is basically a trial run that sets learning and assessment expectations and provides an opportunity for feedback from the instructor. The stage begins with an overview of the method and a logistical walk-through and practical demonstration of the discussion board. Student questions are answered, and then student groups of about five begin their initial 12-day discussion with the expectation that each student will contribute five or six substantive posts. During this first stage students are continually reminded of the aims and expectations of the activity and if there are misunderstandings, poor participation, or students are off-task, the instructor will provide additional guidance within the discussion board itself. However, the discussion should be student led because teacher presence can hamper the validity of an assessment, so every effort is made to keep instructor participation to a minimum. Students need to consider the scenario, the prompts, and the posts of others to generate responses and solutions. At the end of the 12 days, some of the postings are anonymized and shared to serve as a learning resource for the students. Using the posts as learning resources, the instructor leads a discussion about high and low quality responses and how weak posts or discussions might be altered to better meet expectations. The second stage is nearly a mirror of the first except that students utilize a different scenario and they have more independence from the instructor. The transcripts from the second stage are the ones formally assessed using the CPSA rubric.

4.5 Instructors' Manual Exemplars

Issues have emerged organically as the rubric has been used to grade scores. Because of this, we have continually revised the definitions, criteria, and descriptors, all of which has led to inclusion of exemplar transcript excerpts in the administration manual (see https://www.cpsa.ae/wp-content/uploads/CPSA-Manual-11.9.2020.pdf). The role of the exemplars is to show and to explain to raters how a specific rubric score emerges from the transcripts. For example, an exemplar for CPSA 2 is presented below.

Excerpts 2.1 and 2.2
Excerpt 2.1 is representative of a transcript scored in the 1-(Emerging) range as it notices another student's idea as a simple agreement, while Excerpt 2.2 is very similar in terms of agreement and adding their own opinion. It is more sophisticated because the author attempts to bring in colleagues to the discussion, so it is rated as 2- (Developing).

2.1 I agree with you Hajer about what you discussed and I think that. . .
2.2 I agree Salha, I think what companies are doing when collecting information about us and using it the way the want is totally unethical, because it is considered an infringement of the person's rights. I was wondering if Noor or Shama. . .

5. Method

Besides examining the degree to which students are attaining the expected level of proficiency in the professional skills, this research also examines the reliability and validity of CPSA scores within this unique context and specifically for English as a second language learners. The level of student proficiency is measured through the scores assigned by faculty raters using basic descriptive statistics. The reliability mechanisms utilized were the norming process and calculations of interrater reliability because a reliable instrument is one that produces the same scores consistently [29]. The construct of validity is presented as an "argument for what are appropriate interpretations of scores, uses, and consequences, as well as their rationale and evidence that supports that argument of use, interpretation, and consequences" [11, p. 111]. This goes beyond the more traditional and outdated interpretation of validity where there are specific types of validity such as construct, content, and criterion, and positions validity within a particular context.

5.1 Sample

All of the CPSA research has been conducted within a computing program at a public university located in the United Arab Emirates. The program offers English-medium, full-time bachelors degrees and part-time masters degrees that targets working computing professionals. The institution is accredited by the Middle States Commission on Higher Education, one of six United States-based regional accreditors, and the computing programs are accredited by ABET. Students in the programs are nearly all native Arabic-speakers who have studied English as part of their school experiences. Students are eligible to study for a bachelor's degree with an equivalent International English Language Testing System (IELTS) score of approximately 5.0 and at the master's level a 6.0. The English proficiency levels of undergraduates could be described as intermediate, while for graduate students as advanced. From this student population, two groups of 4–5 students were randomly selected from each of the three third year computing classes for a total of six groups. In total there were 25 participants.

5.2 Student Proficiency

Student proficiency is measured through the scores assigned by faculty to each of the groups' transcripts. This is the mean scores of student groups, the mean scores of each of the professional skills learning outcomes and in both cases the percentage of the assigned scores at or above the expected target for third year students of 3.0.

5.3 Reliability

Reliability of an assessment instrument is of the utmost importance because it is a precursor to any arguments put forth concerning the validity of an instrument. The reality is that an instrument can be reliable yet lack validity, but it cannot be valid if it lacks reliability. The norming process and calculations of interrater reliability are the two mechanisms that have been used to measure reliability of the CPSA.

Norming, also known as moderating or calibrating, is a collaborative process where faculty implementing a shared rubric use both knowledge of the rubric and the constructs it represents along with evidence-based discussions to establish a shared understanding of the criteria, descriptors, and levels of attainment present within a rubric. It is a vital procedure because it "does reduce variability across graders" [30, p. 331], and makes the implementation of a rubric more reliable. For the CPSA, a nine-step process for norming has been developed. The steps in the process are: (1) document preparation; (2) rubric review; (3) initial reading and scoring of one learning outcome; (4) initial sharing/recording of results; (5) initial consensus development and adjusting of results; (6) initial reading and scoring of remaining learning outcomes; (7) reading and scoring of remaining transcripts; (8) sharing/recording results; (9) development of consensus and adjusting of results [31].

While the norming process helps to establish consistency between the raters, periodic checks on interrater reliability also serve as a read on instrument reliability for a given implementation and context. Given the relatively small sample sizes and that this is a group assessment, interrater reliability has been calculated using the simple method of percentage agreement amongst raters. This was calculated by dividing the number of transcripts with matching scores by the total number of transcripts. For each implementation of the CPSA where interrater reliability has been determined, this was done overall and for each of the six professional skills measured by the rubric. While the goal for interrater reliability has always been to be as close to 100% as possible, especially considering our extensive norming work, our working threshold for CPSA interrater reliability has been Stemler's [32] 70% target.

5.4 Validity

Because the CPSA purports to be a quality instrument, the validity of the scores is based upon a process whereby multiple sources of evidence have been collected, and from that a coherent argument emerges for the instrument's use [11]. This unitary understanding of validity is well-aligned with the American Psychological Association's definition of validity as "the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of conclusions drawn from some form of assessment" [33]. Evidence to argue for CPSA validity comes from the use of an asynchronous discussion board, language considerations to ensure scenarios are set to appropriate levels for this English as a second language environment, well-constructed definitions of the learning outcomes, and iterative rubric development.

6. Results

Given the three unique aspects of this research, results will be presented specific to student proficiency, instrument reliability, and instrument validity. Student proficiency in the professional skills will be examined by group and by learning outcome (see Table 3). For the groups, the overall mean score was 2.79 while the range of scores was 2.17–3.44. With a target of 3.0 or higher for 3rd year students, only two (groups 3 and 6) of the six groups achieved the target mean. However, when focussing on the percentage of CPSA outcomes $\geq$ 3 for each group, only group 1 was below 50% while both groups 3 and 4 had 83.33% of the outcomes at or above the target, and group 6 was a perfect 100%. From the perspective of the outcomes, only *CPSA 4*, *students communicate professionally in writing*, was $\geq$ the target of 3.0 with a 3.33, yet all outcomes but *CPSA 5. Students analyze the consequences of existing computing solutions within local and global context* and CPSA 6, *students interpret, represent, and seek information*, were at least at 66.67%. This indicates a clear weakness towards the two professional skills.

Results for the reliability of the CPSA are based upon the norming process and findings from the interrater reliability measure. It appears that the nine step norming process has helped graders to produce consistent and reliable scores. The key aspects to this process are that faculty graders

Table 3. CPSA rubric scores

Group	CPSA 1	CPSA 2	CPSA 3	CPSA 4	CPSA 5	CPSA 6	Mean	Group % $\geq$ 3
1	3	2	2	3	2	1	2.17	33.33
2	2	3	3	4	2	2	2.44	50
3	3	2	3	4	3	4	3.17	83.33
4	3	3	3	3	3	2	2.83	83.33
5	3	3	3	3	2	2	2.67	66.67
6	4	4	3	3	3	4	3.44	100
Mean	2.89	2.56	2.89	3.33	2.61	2.44	2.79	69.44
Outcome % $\geq$ 3	83.33	66.67	83.33	100	50	33.33		

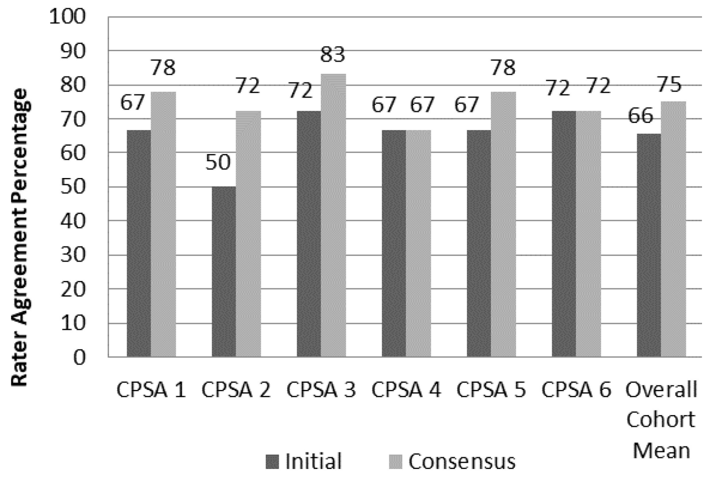

Fig. 1. Interrater reliability.

discuss scores and come to consensus by sharing examples from the discussion transcripts. Faculty graders also utilize the administration manual to examine exemplars to better gauge the scores that they have assigned. These evidence-based discussions have proved to be effective and support for this emerges through the interrater reliability results.

The interrater reliability measure used was the simple method of percentage agreement amongst raters. Results from Fig. 1 indicate that after faculty participated in evidence-based discussions to achieve consensus, this was accomplished for all learning outcomes to the 70% target other than with CPSA 4, *students communicate professionally in writing*, which remained at the same level of 67%. CPSA 3. *Students consider ethical, legal, and security aspects* had the highest rater agreement for initial rating (72%) and again after consensus (83%). Levels of interrater reliability post-consensus ranged from 67–83%, and the overall mean was 75%.

Though we recognize that determining validity is an ongoing process, at this stage the suitability of an asynchronous discussion board, scenarios set to an appropriate English level, the professional skills as clearly defined constructs, and iterative rubric construction, combine to put forth a strong argument that use of the CPSA is valid within this context.

Though its foundation is from the EPSA, a face-to-face instrument, the CPSA is completed online because of the suitability of this medium in such tasks. From a researcher's perspective, the immediate availability of accurate transcripts simplifies the ratings process, and the impact on students' ability to demonstrate learning is in evidence. In addition, we have found that most students talk more in an online group discussion than in a face-to-face classroom. For example, with this implementation of the CPSA, there were a total of 116 significant student contributions to the discussions, while each student posted at least four times. The online group discussion seems to give the shy, quiet, or more passive learner the opportunity to demonstrate their knowledge that can be missed in a traditional face-to-face discussion.

Another way in which the validity of the CPSA is demonstrated is through the appropriateness of an asynchronous discussion board in a second language environment. To achieve this, each scenario is kept to around 700 words with a level of language complexity aligned with the abilities of moderately advanced second language undergraduate students. This means that readability is set to grade 12 on the Flesch-Kincaid scale [34], a suitable level for non-native English language speakers' abilities.

In terms of the professional skills as clearly defined constructs, we believe this has been achieved through the development of the CPSA learning outcomes. By taking ABET's professional skills and creating our own tightly aligned CPSA learning outcomes, defining them further, and then using this to help guide the iterative rubric construction, we have enhanced the validity of CPSA scores as a measure of professional skills within this context.

Iterative construction of the CPSA rubric contributes to its validity within this context. This has been achieved through an initial 2013 pilot implementation followed by multiple iterations with substantial revisions that has produced the current 2020 rubric. Examples of revisions that have occurred include: (1) the first rubric contained only 3 descriptors for each assessment criteria, while later versions have four; (2) an earlier version contained only two assessment criteria (Problem & Solution Identification and Stakeholder Perspectives) for CPSA 1, while the latest versions have

Table 4. Iterations of the CPSA rubric

Spring 2013	Pilot run with one class of 3rd year students using 4 ABET computing professional skills – 24 students.
Spring 2014	Two classes of 3rd year students using 5 ABET professional skills – 42 students.
Fall 2014	One class of 3rd year students using the first complete CPSA with 6 ABET computing professional skills – 31 students.
Spring 2015	Two classes of 3rd year using a revised CPSA with 6 ABET computing professional skills – 38 students.
Fall 2015	Three classes of 3rd year students using a further revised CPSA – 71 students.
Spring 2016	One class of postgraduate students using a further revised CPSA – 17 students.
Fall 2016	Three classes of 3rd year students– 58 students.
Spring 2017	Seven classes of 3rd year students (123 students) and two classes of postgraduate students (34 students).
Fall 2017	Three 4th year classes with 48 students.
Spring 2018	Three 3rd year classes with a total of 41 students, two 4th year classes (43 students), and one graduate class with 14 students all using an updated rubric.
Fall 2018	One 2nd year class (24 students), one 3rd year class (29 students), one 4th year class of 9 students.
Spring 2019	Three 2nd year classes (67 students), two 3rd year classes (51 students), one 4th year class (4 students), and two graduate classes (8 students).
Fall 2019	Four classes of 3rd year students (87 students) and one 4th year class (24 students) using an amended rubric.
Spring 2020	Three 3rd year classes (66 students) and one graduate class of 7 students.

three criteria (Problem Identification, Recommendations for Solutions, and Stakeholder Perspectives). Up until now, 11 faculty and 984 students from 2nd year to masters level have used the various versions of the CPSA. The implementations and iterations of the CPSA rubric are shown in Table 4.

7. Discussion

The main findings of the study are that most students are attaining the desired level of proficiency in the professional skills learning outcomes and that the CPSA has proven to be a reliable and valid instrument to measure the professional skills within this specific context. With its theoretical grounding in learning theory and specifically the COI model [23], the CPSA has shown that it facilitates assessment of complex tasks, demanding teamwork, critical thinking and problem solving, that goes far beyond traditional assessments focused on correct or incorrect responses to discrete, out-of-context items. The interplay between social and cognitive presence demanded by the online asynchronous discussion board task is well-aligned with the reality of the professional skills and offers students an opportunity to demonstrate their proficiency in these cross-disciplinary learning outcomes. The CPSA method has many features that allow all of the skills to be effectively demonstrated by the students. The 12-day period for the activity provides time for reflective thinking, an important component of problem solving, and it provides time for students to clearly articulate their ideas in writing. The discussion board allows students to work in their own time and space and it facilitates problem solving. The team, in the virtual environment, must work closely together on the problem and thus readily demonstrate collaboration, diplomacy, conflict resolution and leadership.

Concerning students' proficiency in the professional skills learning outcomes, it appears that most 3rd year students are performing at the desired level even with some apparent shortcomings. The professional skills *CPSA 5. Students analyze the consequences of existing computing solutions within local and global context* and CPSA 6, *students interpret, represent, and seek information*, were areas of weakness. This mirrors earlier research into student proficiency profiles that found these outcomes low for 3rd year, 4th year, and masters students [35]. These are two outcomes that should be prioritized within the program as they remain a programmatic shortcoming.

Through this implementation of the CPSA, reliability through the measure of interrater reliability had an overall mean of 75% which is within the range of earlier research. For example, a 2019 paper [7] that examined CPSA interrater reliability across 5 separate implementations over a 5-year period found that the interrater reliability through percent agreement had an overall mean range of 74–91%. Though this current implementation was near the lower end of agreement, it was still above Stemler's [32] 70% threshold. Reliability through the norming of the CPSA rubric has been detailed in earlier research [36] and continues to be a strength of the CPSA. Building upon work by Holmes and Oakleaf [37] and Crisp [38] where they provided rules and steps for norming, we have integrated more consensus building dialogue into our process to ensure faculty raters feel listened to, and that a true consensus has emerged within the ratings process.

The case for the validity of the CPSA has been presented through numerous means. A major benefit of using an asynchronous discussion board is that it gives students more time to discuss and solve the problem posed in the scenario which can help foster critical thinking [23]. Based upon student performance on this implementation of the CPSA, we believe that this is the case. In addition, Salmon [24] found that knowledge construction and team work also develop significantly in asynchronous discussion. Given that knowledge construction is closely related to problem-solving, and teamwork was not an area of concern in this study, this has again been supported. Asynchronous discussion is also suitable for most second language students because of the asynchronous nature of these discussions. Krashen's [39] monitor model and affective filter hypotheses can be used to offer an explanation as to why asynchronous discussion works in this way. Krashen posits that with more time students can edit or monitor their output and with this time may have the necessary confidence to communicate more effectively. As they are more confident, they are more willing to communicate and contribute to discussion. With the scenarios set to an appropriate English level of a grade 12 on the Flesch-Kincaid scale [34], the validity of the instrument is improved because it is not an assessment of reading, it remains a test of the professional skills.

8. Conclusion

The CPSA was developed with the aim of creating a rigorous and robust method of assessing all six ABET professional skills for computing within a single assessment instrument that would be suitable for ESL students studying in an English medium computing program. Framed around the constructs of cognitive and social presence, the design of the CPSA method facilitates a rigorous assessment of all the professional skills. Over a period of seven years, it has been improved through a number of iterations and has proven to be a reliable and valid instrument within this specific context. The study found that most students are attaining the desired level of proficiency in the professional skills learning outcomes through the use of the CPSA method.

References

1. Hart Research Associates, Falling short? College learning and career success, 2015, https://www.aacu.org/sites/default/files/files/LEAP/2015employerstudentsurvey.pdf., Accessed 10 January 2018.
2. M. A. Mardis, J. Ma, F.R. Jones, C. R. Ambavarapu, H. M. Kelleher, L. I. Spears and C. R. Mcclure, Assessing alignment between information technology educational opportunities, professional requirements, and industry demands, *Education and Information Technologies*, **23**(4), pp. 1547–1584, 2018.
3. Organization for Economic Cooperation and Development, Better skills, better jobs, better lives: A strategic approach to education and skills policies for the United Arab Emirates, 2015, Accessed 3 November 2017, http://www.oecd.org/countries/unitedarabemirates/A-Strategic-Approach-to-Education-and%20Skills-Policies-for-the-United-Arab-Emirates.pdf
4. H. Jang, Identifying 21st Century STEM Competencies Using Workplace Data, Journal of Science Education and Technology, **25**(2), pp. 284–301, Nov. 2015.
5. P. Griffin, B. McGaw and E. Care, (Eds.), Assessment and teaching of 21st century skills, Springer, London, 2012.
6. L. J. Shuman, M. Besterfield-Sacre, and J. Mcgourty, The ABET 'Professional Skills' – Can They Be Taught? Can They Be Assessed?, *Journal of Engineering Education*, **94**(1), pp. 41–55, 2005.
7. M. Danaher, K. Schoepp and A. Ater Kranov, Teaching and measuring the professional skills of information technology students using a learning oriented assessment task, *International Journal of Engineering Education*, **35**(3), pp. 795–805, 2019.
8. L. Lattuca, P. Terenzini, D. Knight and H. K. Ro, 2014, 2020 Vision: Progress in Preparing the Engineer of the Future Accessed 10 January 2017, https://deepblue.lib.umich.edu/bitstream/handle/2027.42/107462/2020%20Vision%20FINAL.pdf?sequence=5&isAllowed=y.
9. A. Colby and W. M. Sullivan, Ethics Teaching in Undergraduate Engineering Education, *Journal of Engineering Education*, **97**(3), pp. 327–338, 2008.
10. R. M. Felder, 2012. Engineering Education: A Tale of Two Paradigms, in B. McCabe, M. Pantazidou and D. Phillips(eds), *Shaking the Foundations of Geo-Engineering Education*, CRC Press, New York, pp. 9–14.
11. K. A. Douglas and S. Purzer, Validity: Meaning and Relevancy in Assessment for Engineering Education Research, *Journal of Engineering Education*, **104**(2), pp. 108–118, 2015.
12. A. Al-Bahi, M. Taha and N. Turkmen, Teaching and Assessing Engineering Professional Skills, *International Journal of Engineering Pedagogy*, **3**(3), pp. 13–20, 2013.
13. D. Lopez, J. L. Cruz, F. Sanchez and A. Fernandez, A Take-Home Exam to Assess Professional Skills, *2011 Frontiers in Education Conference*, Rapid City, SD, Oct. 12–15, 2011.
14. A. D. Christy, Student portfolios for assessing ABET a-k outcomes, *Proceedings of the 2013 ASEE North Central Section Conference, 2013*, https://www.researchgate.net/publication/259972436_Student_portfolios_for_assessing_ABET_a-k_outcomes.
15. S. Richerson, K. McAteer, M. Spencer and S. Scheibler, *A Portfolio Approach to Learning Professional Skills*, *Frontiers In Education Conference-Global Engineering: Knowledge Without Borders, Opportunities Without Passports*, 2007.
16. A. Ater Kranov, C. Hauser, R. G. Olsen and L. Girardeau, A Direct Method for Teaching and Assessing Professional Skills in Engineering Programs, *Proceedings from the American Society for Engineering Education Annual Conference and Exposition*, Pittsburgh, PA, 2008.
17. A. C. Brualdi, Implementing Performance Assessment in the Classroom, *Practical Assessment, Research & Evaluation*, **6**(2), 1998.

18. R. Johnson, J. Penny and B. Gordon, *Assessing Performance: Developing, Scoring, and Validating Performance Tasks*, Guilford Publications, New York, 2019.
19. A. Ater Kranov, M. Zhang, S. W. Beyerlein, J. McCormack, P. D. Pedro and E. R. Schmeckpeper, A direct method for teaching and measuring engineering professional skills: A validity study. *2011 American Society for Engineering Education Annual Conference, Vancouver*, Canada, 2011.
20. L. Suskie, Using assessment results to inform teaching practice and promote lasting learning, in G. Joughin (ed), *Assessment, Learning and Judgement in Higher Education*, Springer, Dordrecht, Netherlands, pp. 1–20, 2009.
21. D. R. Garrison and N. D. Vaughan, *Blended Learning in Higher Education: Framework, Principles, and Guidelines*, John Wiley & Sons, San Francisco, 2008.
22. H. Zhou, A Systematic Review of Empirical Studies on Participants' Interactions in Internet-mediated Discussion Boards as a Course Component in Formal Higher Education Settings, *Online Learning Journal* **19**(3), 2015.
23. T. Anderson, W. Archer and D.R. Garrison, Critical Inquiry in a Text-based Environment: Computer Conferencing in Higher Education Model, *The Internet and Higher Education* **2**(2–3), pp. 87–105, 2000.
24. G. Salmon, *E-moderating: The key to Teaching and Learning Online*, Routledge, New York, 2011.
25. National Research Council, *Knowing What Students Know: The Science and Design of Educational Assessment*, The National Academies Press, Washington, DC, 2001, https://doi.org/10.17226/10019.
26. The Community of Inquiry, https://coi.athabascau.ca/, Accessed 17 January 2017.
27. R. L. Linn and N. E. Gronlund, *Measurement and Assessment in Teaching*, Prentice Hall, Upper Saddle River, NJ, 2000.
28. M. Danaher, A. Rhodes and K. Schoepp, Reliability and Validity of the Computing Professional Skills Assessment, *Global Journal of Engineering Education*, **21**(3), pp. 214–220, 2019.
29. 29. B. M. Moskal and J. A. Leydens, Scoring Rubric Development: Validity and Reliability, *Practical Assessment, Research & Evaluation*, **7**(10), 2000, http://PAREonline.net/getvn.asp?v=7&n=10, Accessed 12 January 2017.
30. B. O'Connell, P. De Lange, M. Freeman, P. Hancock, A. Abraham, B. Howieson and K. Watty, Does Calibration Reduce Variability in the Assessment of Accounting Learning Outcomes?, *Assessment & Evaluation in Higher Education*, **41**(3), pp. 331–349, 2016.
31. K. Schoepp, M. Danaher and A. A. Kranov, An effective rubric norming process. *Practical Assessment, Research and Evaluation (PARE)*, **23**(11), https://pareonline.net/getvn.asp?v=23&n=11, Accessed 2 January 2019.
32. S. E. Stemler, A Comparison of Consensus, Consistency, and Measurement Approaches to Estimating Interrater Reliability, *Practical Assessment, Research & Evaluation* **9**(4), 2004, http://pareonline.net/getvn.asp?v=9&n=4, Accessed 22 June 2016.
33. American Psychological Association Dictionary of Psychology, *Validity*, Accessed 15 January 2018, https://dictionary.apa.org/validity.
34. J. P. Kincaid, R. P. Fishburne Jr., R. L. Rogers and B. S. Chissom, *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) For Navy Enlisted Personnel* (No. RBR-8-75), Naval Technical Training Command Millington TN Research Branch, 1975, Accessed 15 January 2016, http://www.dtic.mil/get-tr-doc/pdf?AD=ADA006655.
35. M. Danaher, K. Schoepp, A. Rhodes and T. Cammidge, *Student Proficiency Profiles through the Computing Professional Skills Assessment*, IEEE EDUCON 2019, April 8–11, Dubai, UAE., Accessed 22 June 2019, https://ieeexplore.ieee.org/document/8725225.
36. K. Schoepp, M. Danaher and A. A. Kranov, An effective rubric norming process, *Practical Assessment, Research and Evaluation (PARE)*, **23**(11), 2018, Accessed 21 June 2019, https://pareonline.net/getvn.asp?v=23&n=11.
37. C. Holmes and M. Oakleaf, The official (and unofficial) rules for norming rubrics successfully, *The Journal of Academic Librarianship*, **39**, pp. 599–602, 2013.
38. E. A. Crisp, Calibration: Are you seeing what I's seeing? *Intersection*, Winter, **1**(3), pp. 7–13, 2017.
39. S. Krashen, *The Input Hypothesis: Issues and Implications*, Longman, London, 1982.

Maurice Danaher received his PhD in Information Systems in 2003 from Swinburne Institute of Technology, Melbourne, Australia. He received a Masters degree in Engineering Science from the National University of Ireland, Cork, Ireland in 1981. He was awarded a Bachelor degree in Civil Engineering with Honours by the National University of Ireland, Cork, Ireland in 1977. Currently he holds the position of Associate Professor in the College of Technological Innovation, at Zayed University, Abu Dhabi, United Arab Emirates. He commenced his career as a Civil Engineer in Ireland and subsequently worked as an Engineer with the Boeing Company in the USA and Canada. He later taught Civil Engineering at Singapore Polytechnic, Singapore for three years and then for many years lectured and conducted research in Information Technology at Edith Cowan University in Perth Australia. In 2006 he moved to the United Arab Emirates and worked as an Assistant Professor in Information Technology at Abu Dhabi university and from 2008 at Zayed University and was later promoted to his current position of Associate Professor. His research interests are in Information Technology and Quality Assurance in IT Education. He has published his work in these areas and has over 50 refereed publications. Dr. Danaher is a speaker for the Society of Engineers, UAE and is on the list of speakers for IEEE UAE. He promotes the IEEE student branch at Zayed University. He has won a number of grants in the UAE totaling over US$300,000 for research into quality assurance of IT education.

Anthony Rhodes holds a PhD in Information Security from Queensland University of Technology (QUT), Australia. He is currently employed as Chair, Department of Computing and Applied Technology, College of Technological Innovation, Zayed University, United Arab Emirates. In addition to his research interests in information security (security management, access control and database security), Anthony is an active researcher in the domain of teaching and learning as it relates to the attainment of professional/employment skills required by undergraduates as they make the transition from higher education to the workplace. Dr. Rhodes has been at Zayed University since 2004. Prior to this appointment, he was a Senior Lecturer (School of Software Engineering and Data Communications) at Queensland University of Technology, Brisbane, Australia.

Ashley Ater Kranov received her PhD in adult and organizational learning from the University of Idaho in 2006. She has been an adjunct associate professor in the School of Electrical Engineering and Computer Science at Washington State University in the USA since 2009 and is President of the consulting firm Global Professional Skills Assessment based in Baltimore, MD. From 2011 to 2013, Dr. Ater Kranov served as Managing Director of Professional Services at ABET. Dr. Ater Kranov serves as a reviewer for the National Science Foundation, American Society for Engineering Education conference proceedings, and journals such as the Journal of Engineering Education, the Journal of Sustainability in Higher Education and IEEE Transactions on Engineering Education. Her research areas of interest are direct methods to teach and measure professional skills necessary for 21st century engineering and computing workplace success and how to increase gender equity in the technical fields.