IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 15, 2018, accepted October 31, 2018, date of publication November 14, 2018,
date of current version December 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2881282

Comprehensive Insights Into the Criteria of Student Performance in Various Educational Domains

OMAR ZUGHOUL, FAIEZ MOMANI, O. H. ALMASRI, A. A. ZAIDAN, B. B. ZAIDAN, M. A. ALSALEM, O. S. ALBAHRI, A. S. ALBAHRI, AND M. HASHIM
Universiti Pendidikan Sultan Idris, Tanjung Malim 35900, Malaysia

Corresponding author: A. A. Zaidan (aws.alaa@gmail.com)

This work was supported by Universiti Pendidikan Sultan Idris, Malaysia, under Grant 2017-0025-109-01 and 2017-0312-107-01.

ABSTRACT This paper surveyed the literature on the criteria of student performance in various educational domains in order to establish the coherent taxonomy and figure out the gap on this pivotal research area. The significant search for articles focused on: 1) student evaluation; 2) education-related; and 3) criteria and domain. The three reliable databases are the Web of Science, ScienceDirect, and IEEE Xplore. These databases are sufficiently broad to cover education and technical literature. The set comprised 178 articles. The biggest group (170/178) included various evaluation criteria and domains. Most domains (84/170) covered criteria, such as performance and skills. Another group (8/178) utilized surveys and reviews to characterize the evaluation criteria for specific specializations. Since 2012, researchers have investigated the process of student evaluation in several domains. Regardless of class, the challenges that prohibited the full adoption of evaluation criteria were the focus of articles and recommended mitigations. Although research areas on student evaluation differed, the results were evenly pivotal. This paper affirms the existing research and provides the supplemental areas for future work.

INDEX TERMS Student evaluation, student assessment, student appraisal, educational domains.

I. INTRODUCTION

Educational evaluation and measurement have been increasingly enhanced recently. The evaluation of student performance has become a necessary and significant criterion in higher education assessment. Higher education committees consider the quality of higher education from the perspectives of student performance improvement, and these committees give considerable attention to student learning outcomes based on evaluation dimensions [1].

Assessment is the most important factor of the teaching and learning environment and considered as the center of the learning process. Assessment helps professionals in presenting the progress of students and accomplishments and discovering new learning trends. Furthermore, assessment enables educators to obtain feedback via the assessment process [2].

The evaluation of student development covers the improvement of criteria that affect student performance, such as skill development that is measured by various assessment methods. With technology evolution, student evaluation has progressed and trended, but the processes of criterion identification and utilized selection in assessment are still considered major and significant tasks [1]. Despite problems, evaluation and assessment have been expanding. For instance, problems faced by researchers could have resulted from the lack of communication [3] and sample size [4]–[6]. Challenges in problem-solving are significant but might not provide real outcomes for evaluation [7]. Other concerns include limited number of people responsible for assessment [8], challenges with data [9], evaluation types or tools [10] and criteria [11], [12]. This article aims to survey studies and researcher efforts on student evaluation conducted by teachers or evaluators on the performance of students. The criteria used in the evaluation/assessment differ from that of other studies based on the category of study and the investigated domain. Assessment criteria are significant in measuring student capability, learning outcome, performance, knowledge and skills, from which evaluators encounter challenges. In addition, the researcher can determine the gaps arising from cross-over domains and evaluation criteria and motivation for investigation. Hence, a taxonomy includes the papers that used the evaluation methods to assess students by evaluators. However, studies that used the student evaluation

of teachers, courses and self-assessment are excluded from the study. The taxonomy surveyed the undergraduates who belong to higher education institutions in any domain. The studies investigated are from 2012 to 2017 and are classified into (1) evaluation criteria and (2) domains. Articles collected from the literature are categorized on the basis of the domain used in those articles. The articles have different approaches and methods in student evaluation that utilize various criteria according to convenient studies. The articles could benefit researchers [8] by improving student performance [13], investigating evaluation and assessment methods [14] and guiding assessment and curriculum development [15]. The main objectives of this study are the following:

1. to identify the taxonomy and classify research studies on student performance in various educational domains and

2. to identify the motivations, challenges and limitations for the evaluation of student performance and provide recommendations to facilitate and improve this pivotal research area.

The research questions formulated for this study are as follows:

Question 1. What is the taxonomy for the evaluation of student performance in various educational domains that have been studied within 2012 to 2017?

Question 2. What is the distribution of articles published on this topic?

Question 3. What are the motivations, limitations, challenges and recommendation(s) regarding this topic?

Question 4. What are the evaluation criteria used in these articles to evaluate student performance?

This article is structured as follows. Section 1 introduces the student evaluation. Section 2 explains the systematic review protocol for student evaluation. Section 3 presents the taxonomy of student evaluation. Section 4 illustrates the statistical results coming from student evaluation articles. Section 5 details the motivations, challenges and recommendations. Section 6 states the limitations authors faced whilst completing this research. Finally, Section 7 concludes this paper.

II. SYSTEMATIC REVIEW PROTOCOL

A. METHOD

"Student evaluation" is the most significant keyword in this article. This term eliminates any student who evaluates teaching processes or courses. In addition, the term excludes students who study from elementary to high school. The scope of this article is limited to English literature but considers all education-related fields containing the universal categories of domains and criteria, such as medical, engineering and computer.

B. INFORMATION SOURCES

In this study, three reliable databases were selected as sources for searching related studies. These databases are (1) SciencesDirec, (2) IEEE Xplore library of engineering and technology for technical articles and (3) Web of Science (WoS). The rationale behind this selection is to cover the domain and criteria literature and present a worldwide view of researcher efforts. Figure 1 exhibits that the final set number of articles is 178, which resulted from full-text reading in which ScienceDirect, WoS and IEEE Xplore account for 84, 58 and 36 articles, respectively.

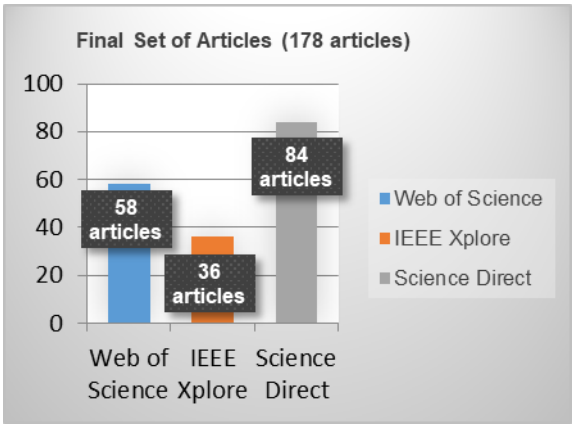

FIGURE 1. Percentage of the final set number of articles in digital libraries.

Two processes were conducted in this article, namely, searching the literature sources and screening and filtering. These processes aim to select articles that are most related to our field. The screening and filtering process consisted of two stages. The first one eliminated the duplicates and unrelated articles by reading the title and abstract of each article. The second stage filtered the articles from the first stage through a full-reading of each article. The authors who conducted these processes used the same eligibility criteria.

The search was conducted at the start of November 2017 in ScienceDirect, IEEE Xplore and WoS databases via their search boxes. We used a mix of keywords that contained "student evaluation," "student assessment," "students benchmark" and "student appraisal" in different variations and combined by the "OR" operator. The exact query text is shown on top of Figure 2. We further used the options in each search engine to exclude book chapters and other types of reports other than journal and conference articles. The years were from 2012 to 2017 in all digital libraries.

The articles were excluded based on the inclusion criteria that we have. Hence, only the articles that mentioned at least one of the inclusion criteria were included and others were excluded.

For the simplification of the filtering process, all included articles were read, analyzed and summarized with their corresponding initial categories and were then saved as an EXCEL® file and Endnote® library. Full text reading for all articles was performed by authors. Numerous highlights and notes on the surveyed works and a running classification of all articles allowed for the creation of the proposed taxonomy. The comments were recorded as hard or soft copy versions

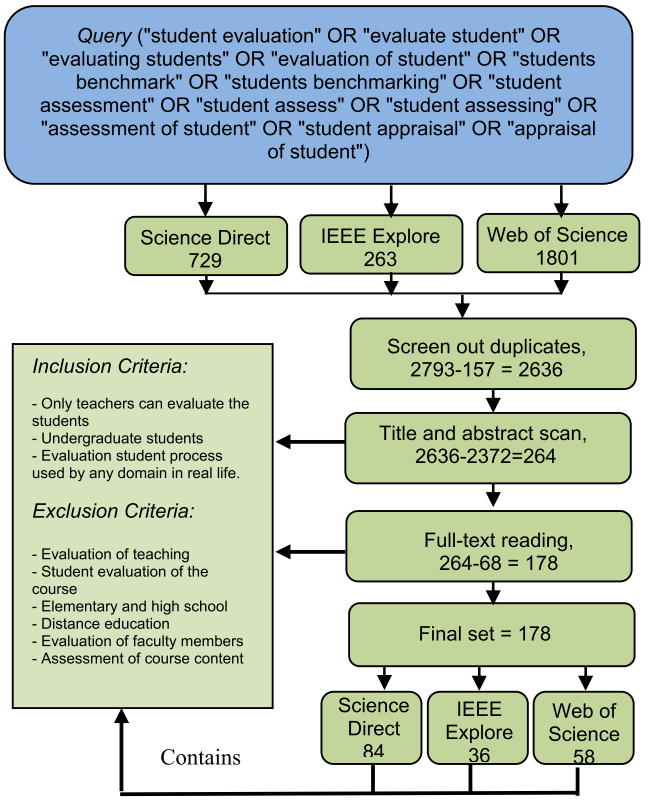

FIGURE 2. Flow chart of the query, selection and inclusion criteria used in the present article.

according to each author's style. This process was followed by another process of the characterisation, description, tabulation and conclusion of essential findings. These findings are provided in the supplemental materials as a complete reference for the results that will be described in the next section.

III. TAXONOMY RESULTS

This study found 2793 articles from 3 databases, which are ScienceDirect, IEEE Xplore and WOS. The breakdown is 729 from ScienceDirect, 263 from IEEE Xplore and 1801from WOS. All these articles are from 2012 to 2017, and the total number of duplicated articles from the 3 databases are 157, which were excluded because of duplication. A total of 2372 articles were eliminated after reading their title and abstract, which further narrowed the number to 264 articles. Finally, after reading each of the 264 articles, we eliminated 86 articles, thereby having 178 articles left in the final list of included articles. To determine the general map for the final set of articles, we conducted a thorough reading to determine the major purpose of these articles. Only a few articles (4.49%; 8/178) are review and survey articles that indicated the evaluation criteria related to various domains. Figure 3 presents a comprehensive overview of this direction.

The largest portion of articles belonged to research articles (95.51%; 170/178). The applied diverse studies ranged from seeking to evaluate student performance from student

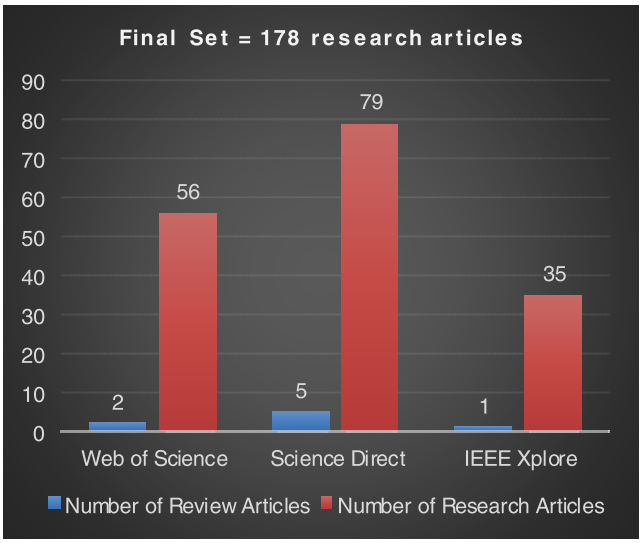

FIGURE 3. Number of review articles compared with research articles.

evaluation to scouting for evaluation-related criteria connected to various domains.

A. REVIEW ARTICLES

The evaluation criteria studied were grouped into various domains. The articles presented the reviews of student evaluation, introduced the criteria to the education community and derived descriptive statistics to understand the implications and possibilities of the research.

The largest group in this category comprised reviews within the Medical domain (5/8 articles). For instance, the review articles conducted in General Medical fields highlighted the criteria used to assess the outcome of student performance, such as grades and effectiveness through summative assessments [16] and [17]. Other articles were used to evaluate student performance through summative assessment and other criteria, such as admission test, GPA and attendance [18]. In [19], formative assessment used by medical institutions, such as using GPA as criterion to evaluate students, was described.

The Nursing field accounted for 1/8 in the Medical domain [20]. This article reviewed the literature on summative assessment of nursing students within nursing and clinical fields.

Another study categorized under the General domain (3/11 articles) used other types of criteria, such as GPA, skills and knowledge, and alternative and innovative approaches that measured student performance evaluations were mentioned in [21]. In [22], which was an assessment of learning outcome articles, academic performances were measured by utilizing the number of credits, using the tests given in a course and the outcome of a progress test using critical thinking, decision making and other criteria to evaluate the 1800 Erasmus students in Netherlands. In [23], in Electrical Engineering (1/8 article), the study conducted a survey of

online quizzes to enhance learning performance. Two groups of students were evaluated in which one group was in a specified course and the other group was in another course. The study was applied in Indonesian universities.

B. RESEARCH ARTICLES

In this section, the taxonomy will be described in detail. Taxonomy includes the number of papers under each domain and the criteria used to evaluate student. In addition, categories and methods used in student evaluation are presented. This study classifies the domains into 12, whereby 11 are specific and 1 is unspecific or called ''general'' because the papers used in this domain did not mention any area of study or they made their studies general evaluation.

1) MEDICAL DOMAIN

The largest rate of articles collected was in the Medical domain (84/170), which consisted of 7 classes. The first class was General Medical and 23 articles were included in this class. According to [24] and [25], medical students should have clinical and communication skills, which are considered important criteria for evaluating medical students. In [5], [26], and [27], four criteria for the evaluation of medical students are stated, namely, ''teamwork skills, communication skills, time management and problem-solving skills.'' References [28]–[31] used GPA as a criterion for evaluating medical students. Reference [32] listed two criteria for evaluating medical students, which are student performance and clinical skills. In [3] and [33], student attitudes are the criteria used. Behnaz *et al.* [34] asserted attitude and knowledge for the students, and in [35] and [36], the criteria for evaluating the medical students are knowledge, skills, teamwork and communication. According to [37], critical thinking is a criterion that measures student performance. Knowledge and skills criteria are used in [38] for student evaluation process. Experiments in [39]–[41] used GPA as the main evaluation criterion. Another experiment study [42] defined performance and skills as criteria for the evaluation tool. A questionnaire is designed in [43] to measure the factors or criteria of 'critical thinking skills'. In [44], the factors were 'knowledge, attitude and performance.'

The second domain is Pharmacy with 16 articles. Performance and teamwork skills are the criteria used [8], whereas knowledge and skills are applied in [45] and [46]. The case study in [47] defined communication skills as the criteria to evaluate students. However, [48] and [49] utilized performance criteria to evaluate Pharmacy students. Student performance and critical thinking skills are the main criteria in [50]. DeGeete *et al.* [51] and Chen *et al.* [52] created a survey to measure student perceptions and O'Brocta and Swigart [9] measured student perceptions and used performance and teamwork skills as criteria. In addition, performance, knowledge, communication, teamwork and critical thinking are the criteria employed in [53] and [54], and Zomorodi *et al.* [55] focused on knowledge and attitude. Knowledge of Pharmacy students is evaluated in [56] and [57], whereas [58] added skills to the knowledge criterion.

The third class is Nursing. The largest rate of articles collected in the medical domain was in this class (33 articles). According to [4], English language proficiency is an important criterion for evaluating nursing students but [13] and [59] defined knowledge, skills and attitudes as criteria for evaluating students. As [60] mentioned, nursing students should improve their clinical judgment. Carter *et al.* [14], Paul *et al.* [15], Chao *et al.* [61], Naber and Wyatt [62], and Shin *et al.* [63] identified critical thinking indicators and its attributes, whereas [64] used ethics as a criterion and clinical judgment skills is employed in [65]. The skills of nursing students are defined in [66] and [67]. However, [68] defined knowledge, skills and critical thinking as the criteria, whereas [69] used only knowledge as criterion. Meanwhile, [70]–[72] mentioned knowledge and skills, whereas teamwork, communication and monitoring skills are cited in [73] and [74]. Performance criteria are mentioned in [75] and communication skills are identified in [76]. Studies [77], [78] defined students' positive attitudes as important criteria for the assessment process. Student performance is the main criterion used in [79]–[82], whereas efficiency criteria are mentioned in [83]. Language spoken at home is a main concern in [84]. Study [85] evaluated students by using pre- and post-simulations, whereas [86] developed a pedagogic learning module, 'Students Active Learning'. In addition, [87] introduced the factors of a new model for evaluating clinical education.

The fourth class is Dental, and five articles comprised this class. In [88], English proficiency is mentioned as important criterion for evaluating dental students. Studies [89], [90] mentioned the performance and skills criteria. In addition, according to [91], critical thinking skills are important criteria for evaluating dental students.

The fifth class is Medicine, and four articles are included in this class. Student knowledge is the main criterion used in [92] and [93]. Academic performance, problem-solving and critical thinking skills composed the criteria for student evaluation in [94]. Hassanien *et al.* [95] developed a CBA system to assess medical students.

The sixth class is Health Care, and three articles are included in this class. Leadership skills and knowledge are the main criteria used in [6] and [96]. In addition, [97] proposed a new version of the 'R-SPQ-2F' for evaluating students.

The last class is Pathology, and this class only has one article. Ho *et al.* [98] evaluated students depending on 'critical thinking, problem solving, communication and teamwork skills'.

2) ENGINEERING DOMAIN

In this domain, the number of articles we found is 21/170, which we divided into 3 classes. The first class is General Engineering, and 13 articles are included in this class. According to [99] and [100], performance criterion is the main factor for evaluating Engineering students.

Knowledge and problem solving skills are mentioned in [101]. Skills and performance are the criteria identified in [102]–[104]. For the other side performance, this criterion is mentioned in [105]–[107]. Performance or GPA criterion is used in [7], [108], and [109]. In addition, 'communication, teamwork, problem solving and critical thinking skills' comprised the criteria in [110].

The second class is Computer Engineering, and three articles are included in this class. 'Project management, teamwork, written communication, presentation and oral communication skills' are defined in [111] as criteria for evaluating Computer Engineering students. In another study, knowledge and skills are used as criteria for evaluation [112], [113].

The last class is Electronics Engineering, and five articles are included in this class. Basic, design, lab, inquiry, profession, communication skills and GPA are the criteria employed in [114], but only GPA is mentioned in [115] and [116]. In addition, knowledge and skills criteria are included in [117] and problem solving skills criteria are stated in [118].

3) CHEMISTRY DOMAIN

In this domain, the number of articles we found is 3/170. Manipulation skills are the main criteria used [119] to assess students, and skills and attitudes comprised the criteria in another study [120]. In addition, performance is the main criterion [121].

4) PHYSIC DOMAIN

In this domain, the number of articles we found is 6/170. Skills criteria are utilized to evaluate students [122], whereas performance is the main criterion in [123] and [124]. Student attitude criterion is mentioned [125], and writing skill is the main criterion in [126] and [127].

5) COMPUTER SCIENCE DOMAIN

In this domain, the number of articles we found is 9/170. In [10], students skills are the criteria for evaluation, whereas programming skills comprised the measure [128]. In addition, proposing a system for evaluating student is conducted in [129] and [130] used student performance for the evaluation process. Student knowledge is an important criterion to evaluate Computer Science students [131], whereas skills criteria are the main concern in [132]. Research [133] aimed to develop a tool for the automated evaluation of programming assignments, such as FSMDA, and [134] proposed a new model for the evaluation process. 'A method to assess POs by adopting the approach of mapping CLOs with POs based on the mapping criteria of multi-academic accreditation bodies' was developed in [135].

6) MARKETING DOMAIN

In this domain, only one paper was found related to our research. Collaborative learning tools using CM is developed [136] to evaluate Marketing students by measuring their knowledge.

7) BUSINESS ADMINISTRATION DOMAIN

In this domain, the number of articles we found is 4/170. In [137], the criterion used for student evaluation is performance, whereas knowledge and skills are utilized as criteria for evaluation in [138]. Similarly, performance criteria are used in [139] and [140].

8) LANDSCAPE DOMAIN

In this domain, the number of articles we found is 2/170. Skills comprised the main criteria used for evaluating the landscape programme in [12] and [141].

9) INDUSTRIAL DESIGN DOMAIN

In this domain, only one paper is related to our research. Salvador *et al.* [142] developed a mechanism that acted as an assistant to the instructor by automatically evaluating students based on the performance criteria.

10) SOCIOLOGY DOMAIN

In this domain, only one paper is found related to our research. In [11], an assessment method is designed to evaluate Sociology students based on knowledge and writing skills criteria.

11) ENGLISH DOMAIN

In this domain, the number of articles we found is 2/170. Performance comprised the main criteria used in [143] to evaluate English students and skills criteria are employed in [144].

12) GENERAL DOMAIN

The second largest number of articles we collected is in the General domain (36/170). In [145], skills composed the main criteria to evaluate students and performance was used in [146] and [149], as well as in [150]–[152]. Performance with skills and attitude are listed as criteria in [153] and [154]. In addition, critical thinking skills and attitude criteria are used in [155], whereas performance and language are applied in [156] and [157]. References [158] and [159] developed an assessment system to evaluate students. A quantitative method is used in [160] and qualitative methodology is utilized in [161]–[163]. Reference [164] developed a methodological approach, whereas [165] proposed a new evaluation method based on fuzzy method. Zhang and Yang [1] proposed a model and applied Support Vector Machine theory to evaluate students. Reference [166] put forward a conceptual model, whereas [167] presented problem-solving criteria as the main criteria for evaluating students. Moreover, student attitude and attendance are the main criteria used in [168] and [169]. In addition, attendance and student final grades are the criteria used in [170] and self-work skills are applied in [171]. Problem solving is the criterion used in [172] and knowledge is the main criterion in [173] and [174]. Moreover, knowledge with performance and skills criteria is mentioned in [175], and knowledge with skills added is in [176]. Learning skills

comprised the main criteria in [2]. Communication, teamwork and task management skills are the criteria reported in [177], whereas knowledge, skills and attitude criteria are mentioned in [178].

IV. DISTRIBUTION RESULTS

The research articles are divided into two categories, namely, (1) evaluation criteria and (2) domains. Articles are divided into two categories to analyze the papers that include a criterion for the evaluation process. The review articles only review the works and not study the criteria that affects student evaluation. The domains are categorized into 12 with 11 specific and 1 general that refers to different unspecified domains. The domains are Medical, Engineering, Chemistry, Physics, Computer Science, Marketing, Business Administration, Landscape, Industrial Design, Sociology, English and General. Figure 4 shows the domains and the classification of these domains. Most of these papers are in the Medical and Engineering domain at 49.41% (84/170) and 12.35% (21/170), respectively. In addition, the General domain has 36 out of 170 studies (21.18%; 38/172), whereas the Computer Science domain has 9 out of 170 studies (5.29%). The Medical domain is subcategorized into seven fields, in which the Nursing field has the largest number of studies (39.29%; 33/84) and Pathology field only has one. The Engineering domain has three fields in which General Engineering has 13, Electronics has 5 and Computer has 3. The least number of studies are in Marketing, Industrial Design and Sociology, as illustrated in Figure 5.

On the opposite side of the taxonomy is the Evaluation Criteria, which is categorized into Performance (GPA), Skills, GPA and Skills, Other Criteria, GPA and Skills and Other Criteria, GPA and Others and Skills and Others. Other criteria are subcategorized into various terms, such as Knowledge, English Proficiency, Attitudes, Attendance, Ethics and Efficiency. Knowledge criteria are studied as a research conducted in the Marketing domain, whereas GPA and Skills as evaluation criteria are studied within different domains and fields (General Medical 23; Electronics 5; Physics 6). Figure 6 presents the categories. We could differentiate subcategories in the major classes, as overlaps happen. The observed categories are listed in the next sections, with simple statistics throughout the discussion. Figure 6 shows categories of studies by year of publication. Data indicated that 2014 had the most studies published in WoS database, whereas ScienceDirect had the most number of studies published in 2016.

Figure 7 presents the number of included articles according to publication year. The distribution of articles from 2012 to 2017 is shown as well.

Furthermore, Figure 8 indicates that authors who published research on student evaluation came from 37 countries. We observed that the studies are directly from certain countries or cover a case study in these countries.

The nationality distribution of the 178 student evaluation papers in frequencies and percentages shows that the most productive authors are from the USA. The 'others' category comprised Australia, Spain, UK, Malaysia, India, Saudi Arabia, Republic of Korea and Taiwan; three articles each are from Czech Republic, Ireland, Pakistan, Hong Kong, Greece and China; two articles each are from The Netherlands, Canada, Sweden, Kuwait, Japan, Iran, Turkey, Romania and South Africa. Only one article each came from the UAE, France, Egypt, Iceland, Thailand, Germany, Portugal, Shanghai, Kazakhstan, Slovakia, Norway and Italy. The reason for analyzing the articles based on countries is to explore the type of research done country-wide.

V. DISCUSSION

This study aims to present updated and state-of-the-art findings on the research about evaluation of student performance to highlight research trends. The present systematic review differs from previous reviews in terms of recentness and focus on literature rather than on domains, evaluation criteria and cross-overs. Furthermore, the study proposes a taxonomy of related literature. Developing the taxonomy of literature in a research area, particularly an emerging one, has several benefits. Firstly, a taxonomy of published works organizes publications. A new researcher who is studying the evaluation of student performance may be overwhelmed by the large number of papers on the subject without any kind of structure and may fail to obtain an overview of this area. Different papers could be reviews or examinations of tools and methods used in student evaluation with important criteria. A taxonomy of the related literature, as shown in Figure 3, could systematize these different studies and activities into a meaningful, manageable and coherent layout. Secondly, the structure introduced by the taxonomy provides researchers with important insights into the subject in several ways. Such structure outlines potential research directions in the field. For example, the student evaluation taxonomy in this work reports researchers are inclined to propose or examine evaluation methods in the medical domain, and thus, focusing on this part can be a prospective research topic. Other research paths include engineering and computer science. A taxonomy can reveal the gaps in certain areas. Mapping the works on student evaluation into distinct categories highlights the weak and strong spots of research coverage. For example, the taxonomy in this article shows considerable attention in the Medical domain (reflected from the proliferation of their classes and subclasses) at the expense of other domains. Statistical data on individual categories of the taxonomy identify the involved sectors in student evaluation to keep up with new trends and strengthen lesser active areas. Similar to taxonomies in other fields, a common language is developed amongst researchers to communicate and discuss emerging works, such as development papers, comparative studies and reviews on student evaluation. Based on the reading of literature content, most of the student evaluation articles include motivations behind student evaluation, challenges to the success of student evaluation and recommendations to alleviate these difficulties.

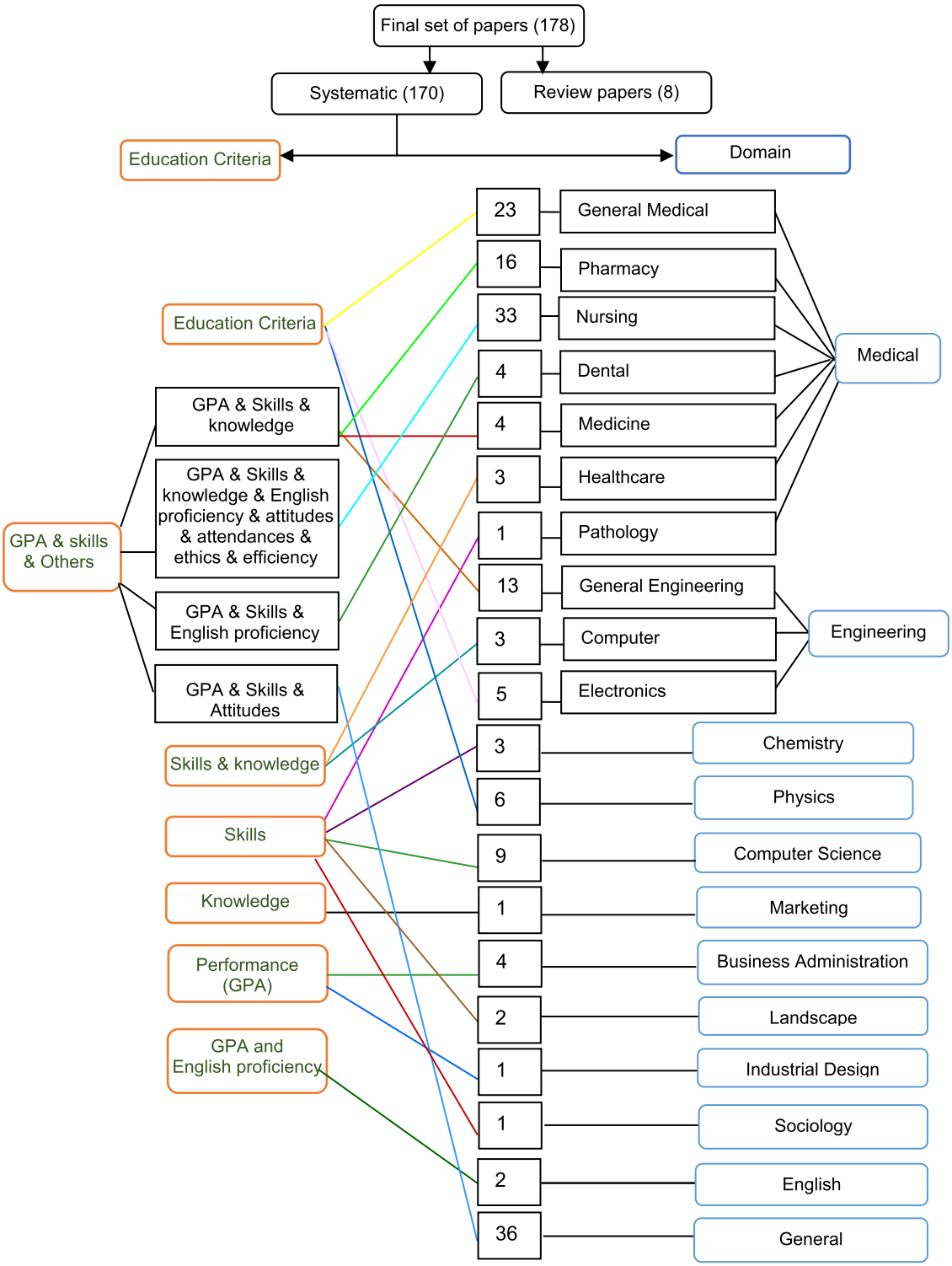

FIGURE 4. Cross-over taxonomy of search literature on student evaluation.

A. MOTIVATIONS

This section presents the motivations discussed in the literature and are arranged as groups of similar benefits. Citations are included for each benefit for further discussion. Figure 9 summarizes the motivations for student evaluation.

1) IMPORTANCE OF STUDENT EVALUATION

The evaluation of student learning achievement is essential for schools, teachers, parents and students [21], [96], [103], [146], [154]. The process for evaluating the performance of students is considered as the most important element in

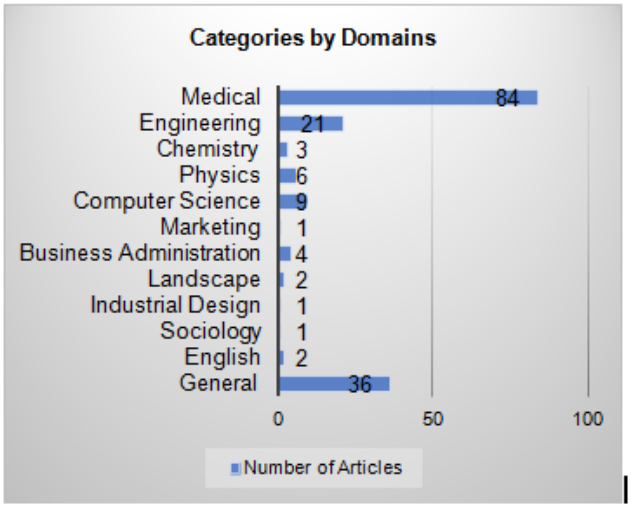

FIGURE 5. Categories by domains.

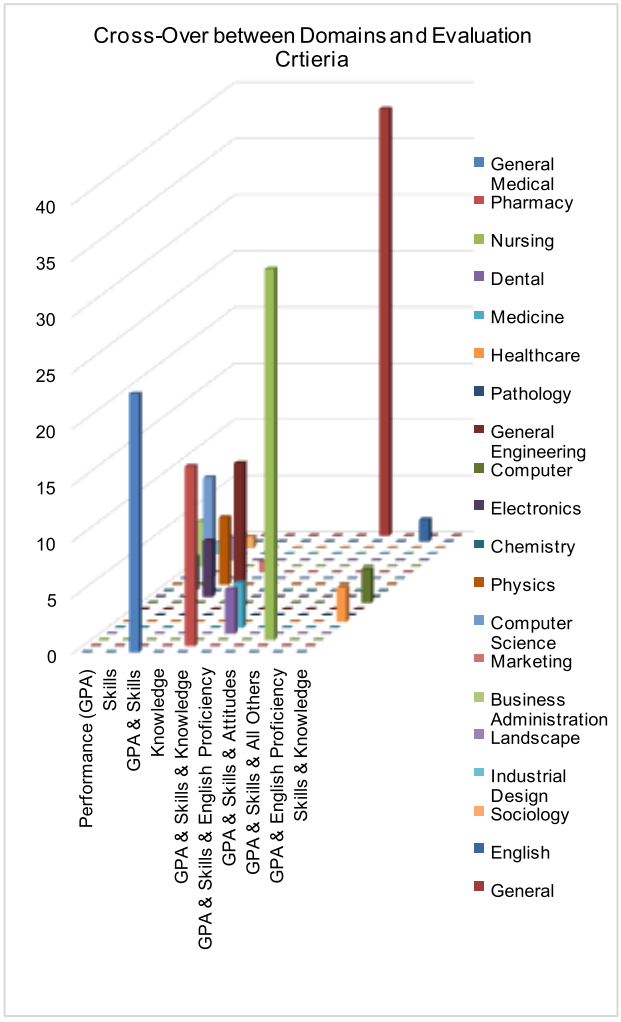

FIGURE 6. Cross-over between the domains and the corresponding evaluation criteria.

the educational system [81], [172]. Thus, the student's skills in handling real-life experiences should be evaluated. The evaluation process helps in identifying the weaknesses of

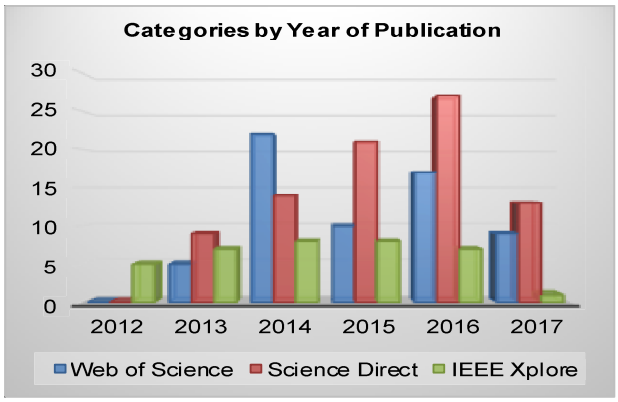

FIGURE 7. Categories by publication year.

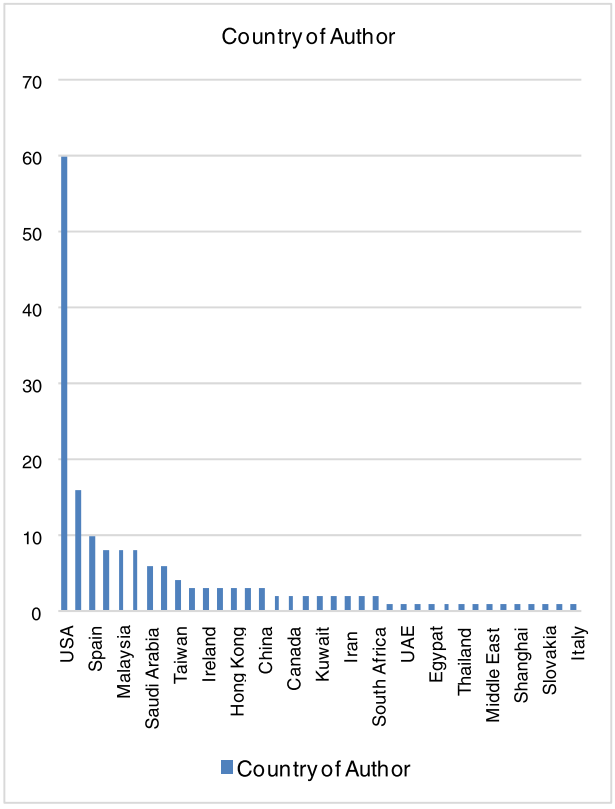

FIGURE 8. Categories by the country of authors.

the students and finding the right way to improve the weaknesses [45]. The evaluation methods of CT at each phase of the nursing process [61] and assessment methods that consider new science education standards are necessary [122]. Predicting student academic performance is usually one of the biggest efforts in EDM [170]. The evaluation of student performance based on curriculum goals is considered significant [119]. Such assessment process plays a major role in enhancing the outcomes of an educational objective programme [135]. Measuring the real knowledge of users is difficult because they can answer in a right manner even

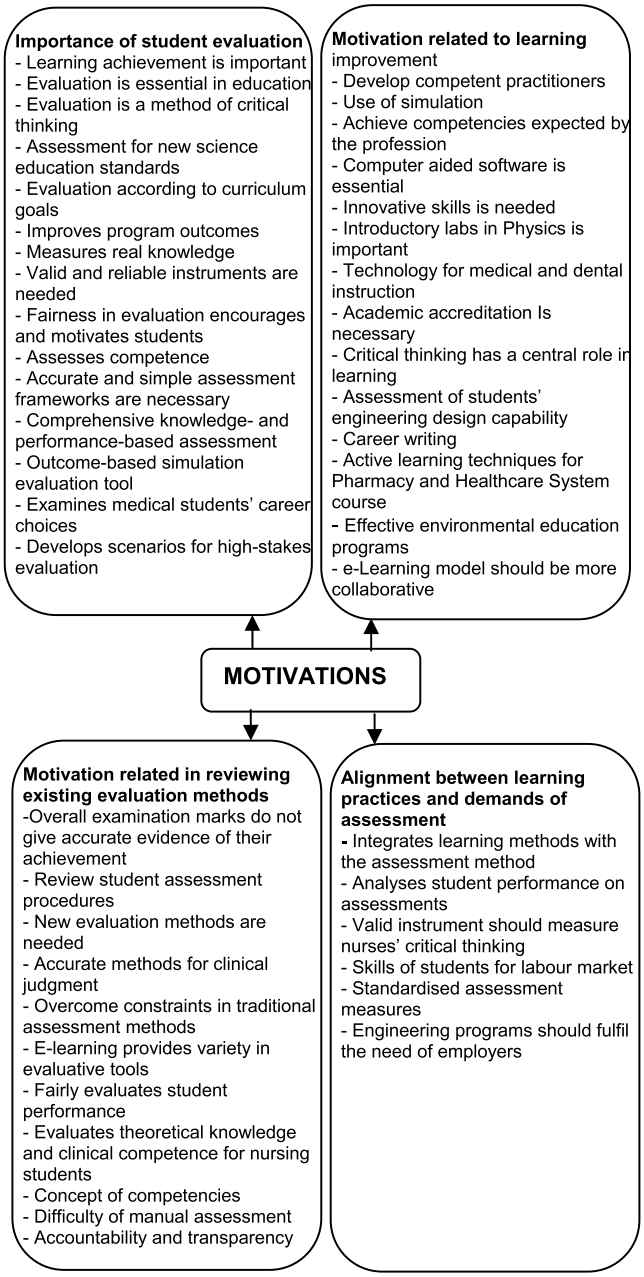

FIGURE 9. **Motivations for student evaluation.**

without gaining formal academic knowledge [10]. Validity instruments are required to allow the objective evaluation of the faculty, provide feedback to each student, and raise teamwork knowledge, skills and attitude [74]. Finding good methods for the assessment of student performance is important for medical schools [27]. Fairness in the evaluation of educational institutions encourages students to perform better. Unfairness in the evaluation process may cause some students to miss career opportunities, which will discourage them from pursuing their goals [165]. The value of competence assessment in practice creates an important requirement in developing responsive policy and process and technical higher education [71], [111]. Thus, the type and value of the assessment and impact assessment on performance in general should be considered [171]. Moreover, schools and colleges of pharmacy should carry out comprehensive knowledge- and performance-based assessments of student achievement in terms of learning outcomes and professional competencies [49]. The development and testing of a healthy communication assessment tool response to physiological changes are equally important [76]. The adaptive tool could help faculties in validating observation results in the clinical setting, guiding the debriefing process and identifying potential areas for curriculum revision [79]. The assessment tool of student performance in PT is an important need in student evaluation from academician point of view [12]. The investigation of the choices of medical student profession, as well as the factors that may influence medical student career performance and choices, is also important [33], [73].

2) MOTIVATION RELATED IN REVIEWING EXISTING EVALUATION METHODS

Evaluating student performance only from overall examination results does not provide accurate evidence of achievement on a particular subject; thus, the use of an evaluation method that accurately provides details is required [123]. References [146] and [162] stated that student assessment procedures should be reviewed in higher education quality assurance. The minimum attendance criterion varied from 60% to 75% and from one university to another [169]. Group studies and classical student evaluation methods are no longer sufficient for qualified education, such that new evaluation methods are needed [112], [118]. For instance, although universities follow a standard regarding class attendance, some universities allow students to not attend 40% of lecturers and if he/she exceeds this limit, he/she will fail in the course. Other universities give different ratios and believe that evaluating students based on the attendance is not enough. In addition, classical student evaluation methods have become insufficient in evaluating the performance of the students. Hence, the needs of new evaluation methods are considered as a motivation for the researchers. Accurate methods for clinical judgment [60] are required to overcome constraints in traditional assessment methods that depend on grades in testing, assignments or quizzes. Such measures lack precision and reliability when measuring the achievement level of each student [124]. In comparison with conventional learning, e-learning provides various evaluative tools [23], [121]. In the same context, [18] confirmed that evaluation systems that are adaptive to electronic issues could create efficient response rates and timeliness with no impact on the quality of responses, compared with other evaluation systems, such as paper-based methods. The quality and ability of education systems are dependent on important issues, such as the assessment of reformation model and scientific improvement of comprehensive evaluation systems [174]. Thus, a model that could help teachers to evaluate student performance fairly is needed [130]. The evaluation of theoretical knowledge and clinical competence for nursing students

are necessary because measuring psychomotor skills and utilisation of theoretical knowledge, judgment and ability to respond to a changing environment are equally important [68]. The adaption of concepts of competence as a rule for learning outcomes that represent a dynamic set of knowledge, understanding, skill and ability accomplished by planned educational processes is required [137]. Assessment programmes for students carried out manually have many drawbacks, such as inconsistency and slower speed than other programmes, particularly when a large number of students and instructors have to accomplish much assessment work manually [133]. Great accountability and transparency, especially in the assessment of student learning, should be ensured in higher education [176].

3) MOTIVATION RELATED TO LEARNING IMPROVEMENT

For the improvement of competent practitioners qualified for delivering high quality nursing care [67], [68], nursing simulation is used to acquire momentum and could control pedagogical methods to mimic a clinical situation [25], [70]. Furthermore, [65] and [82] stated that the simulation is an alternative learning strategy to enhance clinical practice in a safe environment. Students need opportunities to translate their theoretical knowledge into practical applications, be socialized into the profession and achieve competencies expected by the profession. Each student has to acquire appropriate opportunity to translate his/her knowledge into an application with practical and theoretical aspects [25]. Students could be motivated to be an active member of his/her team in the clinical settings and could increase their skills, such as solving problems and CT [15], [98]. Soft skills in using computer-aided software are significant in hand drawing. Reference [141] confirmed that students need to improve innovative skills because they are important in innovation [167]. Introductory lab can be a good avenue for enhancing student attitude and epistemology in physics through experimentation processes [125]. The use of technology should be embedded within medical and dental teaching [90]. Strategies are necessary to encourage dialogue and consensus amongst nurse educators [107]. IT can play a major role in optimisation and streamline classrooms and schools [152]. Academic accreditation is needed to ensure that academic programmes have certain standards essential to encouraging graduate students to involve their professional skills [102], [105]. CT has a primary role in learning and is addressed as a core outcome in higher education [43], [120]. Clinical placement is a central component of undergraduate education in all medical/health science programmes [44]. An important issue in teaching introductory design courses is the assessment of student engineering design capability [104]. Career writing can help students cultivate affective development, improve intrinsic motivation and enhance their ability to make suitable choices [161]. Using active learning techniques has many benefits in the Pharmacy and Healthcare system programmes [9]. Team-based learning improves student performance and perceptions compared with lecture-based courses [53]. A resource-efficient approach for the clinical education of nursing practitioner students is needed [87]. The importance of undergraduate training in tuberculosis was noted in [34]. To implement and evaluate effective environmental education programmes in unique cultural contexts, critical, place-based approaches are evaluated [164]. Many students struggle to understand CT due to lack of confidence in its application and doubt how they could develop CT skills and struggle [69]. The e-learning model should be collaborative with the change in the roles of students and instructors [147]. The need to produce clinically competent pharmacy graduates is essential for sustaining the ongoing effectiveness of the pharmacy profession [58].

4) ALIGNMENT BETWEEN LEARNING PRACTICES AND DEMANDS OF ASSESSMENT

Learning and education methods should be integrated with assessment methods used [94], [110]. Schools and colleges should analyse student performance on assessments for continuous curricular improvement to aid students achieve learning outcomes and professional competencies [49]. CT has been identified as a vital outcome for nursing education. However, the lack of a valid instrument to measure CT abilities of nursing students has limited the assessment of student achievement in Korean academic programmes, leading to ineffective academic mentoring [63]. Thus, student skills should be improved such that learners can easily be integrated into the labour market [176]. Inconsistencies in approaches to clinical teaching and evaluation could result in significant challenges without reliable, standardized assessment measures [80]. All engineering programmes should be developed to fulfill the need of the government and graduate schools and employers should ensure that students have the necessary skills to succeed after graduation [7], [159].

B. CHALLENGES

In recent years, interest in student evaluation has increased, although the area still encounters issues and challenges in many important aspects. Figure 10 illustrates the examples of such challenges.

1) CONCERNS ON LACK OF COMMUNICATION

This section describes challenges related to evaluation difficulty caused by weak communication between evaluators and students. Students may underperform due to low level of English language proficiency [4]. As learners communicate, they are challenged to interpret verbal and nonverbal communication and respond with communication strategies towards desired outcomes [3], [144].

The lack of communication or misunderstanding may cause bad effects, which are considered as challenges. Criteria that affect the process of communication, as mentioned earlier, cited that if the student is weak in the English, he/she might misunderstand issues. Subsequently, the performance of that student will be lower than others. In addition, the communications between the evaluator and the student is

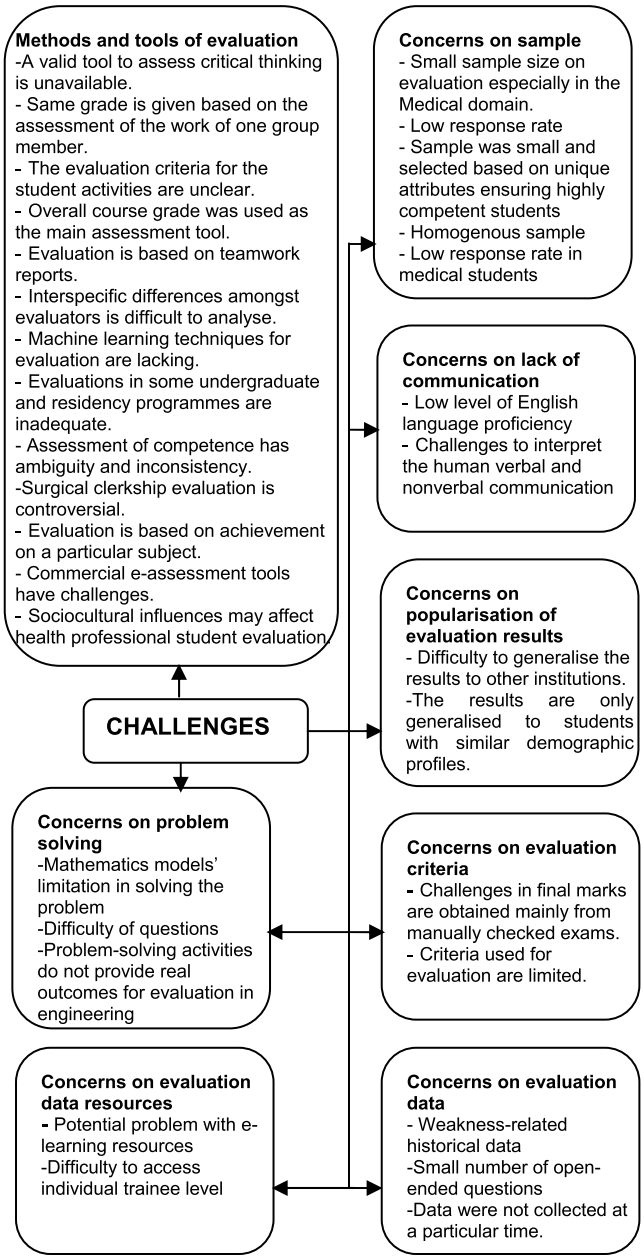

FIGURE 10. **Challenges in student evaluation.**

very important because if the evaluator could not understand what the student says, he/she will give a low score to the student.

2) CONCERNS ON SAMPLE

The challenges related to sampling size and availability of sample in carrying out a research on evaluation are described in detail in this section. A small sample size poses challenges in carrying out research on evaluation especially in the Medical domain. References [4]–[6], [8], [40], [41], [58], [63], and [65] stated that a small sample size is a challenge faced by a group of researchers in completing their study. Reference [9] mentioned that a potential reason for the low response on our study is that not all students in each group have an opportunity to present to the class. Moreover, [51] mentioned that a biased sample could be caused by the low response rate, as students with healthy behaviors may likely complete a health-related survey. The appropriate final conclusions should be drawn from an appropriate sample size because the size of a sample does not reflect reality [126], [161]. The results may not be generalizable to all students because our sample is small and selected based on unique attributes to ensure highly competent students [42], [57]. The challenge is the small homogenous sample size [85]. References [53] and [81] confirmed inconvenience sampling, insufficient sample size and number of students in each comparison group based on a priori power analysis as the challenges in their work. Low response rate amongst medical students despite sending multiple reminders is also a challenge. Thus, sample size might not be ideal to represent the awareness level of all medical students [93]. Reference [67] confirmed that a low number of study participants might have influenced the possibility to achieve statistically significant differences between groups and their goal attainment.

This section describes the limited number of staff available to evaluate students. The staff is unqualified to evaluate students. A large number of students do not match the number of faculty members and staff; thus, the assessment of students is difficult [45]. Limited availability of teaching staff and absence of routine structures associated with traditional classes are amongst the problems faced in assessment activities [66]. Reference [84] mentioned that sole access to a student cohort in one university setting with a limited number of academics teaching and assessing undergraduate nursing students promoted difficulty. Reference [8] confirmed that the topics and faculty facilitators changed slightly in the course each year, which affected the method of assessment and teaching.

3) CONCERNS ON METHODS AND TOOLS OF EVALUATION

This section describes the challenges faced by researchers in using the methods and tools (methods and tools used in evaluating the students were inefficient or not compatible) whilst evaluating students. The traditional evaluation tool results show that when a student obtains high scores, the performance of the other student could be better than him/her even if they get a lower score [10], [50], [172]. A common, reliable and valid tool that could assess CT is unavailable [15]. To encourage fruitful collaboration in the lab, students work in small groups and submit work individually, although group members receive the same grade based on the assessment of the work of one group member [125]. The evaluation criteria for student activities are unclear [8]. Abundant data are collected in each course. However, student understanding of the principles and concepts related to the outcomes is difficult to deduce from the data. The overall course grade is used mostly as the main tool for assessing the achievement of outcomes. Such situation made challenging the identification of areas in which students had difficulties and need continuous improvement [108]. In the evaluation

based on teamwork reports, students are usually unable to report their work properly, and thus, at least several initial reports are misleading [113]. A limitation of this study is that VHI and Voice Evaluation Form are used to assess participant voice and focus on female students of speech and language therapy and use of self-reported questionnaire [157]. Inter-specific differences amongst evaluators or differences within specific rotations is difficult to analyze [28]. Reference [132] confirmed that the lack of machine learning techniques to evaluate the validity of the relationships is included in the concept maps. This result is attributed to the nonexistence of a knowledge base that stores related concepts. No fully formalized design that assigns specific assessors or a specific number of assessors to each portfolio task is available [29]. The amount of clinical work on each service is difficult to objectively quantify [27]. Some undergraduate and residency programmes have inadequate evaluations and often do not allow students to review their assessment [18]. One of the main challenges in using natural language is that no specific fuzzy method is currently available to transform student performance collected entirely as natural language [21]. The quality of an assessment is affected inter alia on assessor's ability to use the assessment instruments provided to make a judgment on performance. Their judgment may be compromised by considering evidence that is not relevant to the competency they are assessing or they may neglect aspects of the performance that are important [44]. A single-point in-time assessment is problematic, and thus, a long-case or clinical practicum examination has been the mainstay of competency assessment in osteopathy [96]. Assessment of competence is characterized by ambiguity and inconsistency despite its critical role in assessing readiness for entry to the nursing profession [71]. Designing quality large-scale assessments is extremely challenging as it is characterized by high stakes of students enrolled in multiple centers and national/state-level competencies for assessments [128]. For example, the evaluation of medical students during surgical clerkship is controversial [28]. The assessment of student performance solely from overall examination scores does not provide accurate evidence of their achievement on a particular subject [123]. Two of the major challenges in commercial e-assessment tools available are that students found the tool interfaces incomprehensible and the administration associated with implementing the tools proved to be arduous [39]. Finding the most effective method for comparing achievement between different cohorts of students is a challenge for practitioners [41]. The evaluation of health professional students through empirical training is considered as subjective assessor judgment and affected by 'sociocultural influences' [92]. Faculty- and house staff-written evaluations for student assessment in medical education, e.g. surgery, have poor specificity and a high degree of subjectivity [32].

4) CONCERNS ON EVALUATION CRITERIA

This section describes the challenges with the criteria used in the studies evaluation, which were inefficient or did not have a procedure to evaluate students. One of the main challenges is the development of suitable assessment criteria [134]. The final marks of checked exams and course work results are gained manually until now and this method influences assessment validity and causes validity problems due to the assessment failures to present scores that could support the valid inference for student writing achievements [11]. The evaluation criteria of the student performance are considered limited because academician evaluators are not frequently with the PT of the students [12].

5) CONCERNS ON EVALUATION OF DATA RESOURCES

This section describes the challenges faced by researchers in choosing the resources of data. A potential problem with e-learning resources is the lack of accountability of learners and difficulty in encouraging the diligent use of resource. During the study period, multiple email reminders are sent to students to encourage the use of the resource; thus, collating data from all students could be collected [90]. We used board certification as a binary measure as opposed to one's certification score and did not include GME level data, such as in-training examination performance or residency performance evaluations because we did not have access to such data at the individual trainee level for this cohort [40].

6) CONCERNS ON POPULARISATION OF EVALUATION RESULTS

This section describes the challenges faced by researchers in generalizing the results of the studies to other cases by using various rationale, such as criteria or method used or ungeneralizable result. The investigation on a single institution may not be generalizable to other institutions [32], [40], [91]. Although most studies are limited by their use of only one school, only students with demographic profiles similar to those reflected in that programme can be considered for generalization [52], [76]. The limitations of using instruments on students in one programme with a sampling approach limits generalizability after the current sample [74]. However, studies conducted in foreign countries, the ScoDoc software systems, should be internationalized [159].

7) CONCERNS ON PROBLEM-SOLVING

This section describes the challenges related to how individuals think and how to solve the problems faced by students with the level of difficulty of exams. The limitations of mathematical models in solving problems related to subjective judgments and variations of human capabilities are mentioned in [150]. More than half of the questions are categorized as difficult, whereas the two most difficult questions could be answered correctly by 33.9% of the students [123]. The problem-solving activities did not provide a real outcome for evaluation in engineering [7].

8) CONCERNS ON EVALUATION DATA

This section describes the challenges related to data and their collection. Prior math coursework and most ALEKS

scores have weaknesses because they are historical data [99]. However, we do not have data to suggest whether this method is adequate for student feedback needs compared with written or typed comments [5]. The small amount of qualitative data was derived from a few open-ended questions [6]. We used board certification as a binary measure as opposed to one's certification score [40] and data are not collected at a particular time. Hence, time of day is not considered a contributing factor and collection of actual clinical assessments is restricted to the staff who conducted the assessment at least for two days and conducted at least 20 assessments across various campuses, which restricted the generalizability of the findings to similar assessment contexts [84], [155]. Although we do not have data to compare student perceptions of a lecture based on top 200 medical courses and a course using a significant amount of active learning techniques, the overall assessment by the students is favorable and student success as defined by final grades is good [9].

C. RECOMMENDATIONS

This section presents important recommendations to the literature. Figure 11 depicts some examples for such recommendations.

1) RECOMMENDATIONS TO IMPROVE STUDENT PERFORMANCE

This section presents the recommendations that contribute in improving student performance. Three studies argued that the use of information technology in educations has positive effect to improve the performance of the students. When applying a combination of online and face-to-face learning in teaching nursing students, they became qualified in administrating IV infusions and gained preparation skills [66]. The students evaluated by using electronic forms and modified completion processes showed high global rating, low missed data rates, high faculty completion rates and consistent collection of common feedbacks, and they had potential to provide a means for others to improve their end-of-shift evaluation completion rate [5]. In the same context, [141] deduced that the use of computer software in education is necessary. Reference [13] mentioned that dependence on the addition of rigorous simulation could improve the performance of the learning of nursing students. Raising the number of assessment activity in each course is beneficial to motivate the use of ICTs and student performance could be enhanced in the final exam if they conduct and finish continuous activities [137]. Reference [112] confirmed the importance of laboratory applications in student performance in engineering education.

References [101] and [136] confirmed that social interaction and collaboration learning have motivating effects on the learning process. RLOs increased the ability and confidence to meet wound care competency outcomes [59]. The sensitivities and adaptabilities of nursing students should be combined as a fundamental quality in undergraduate nursing programme [67].

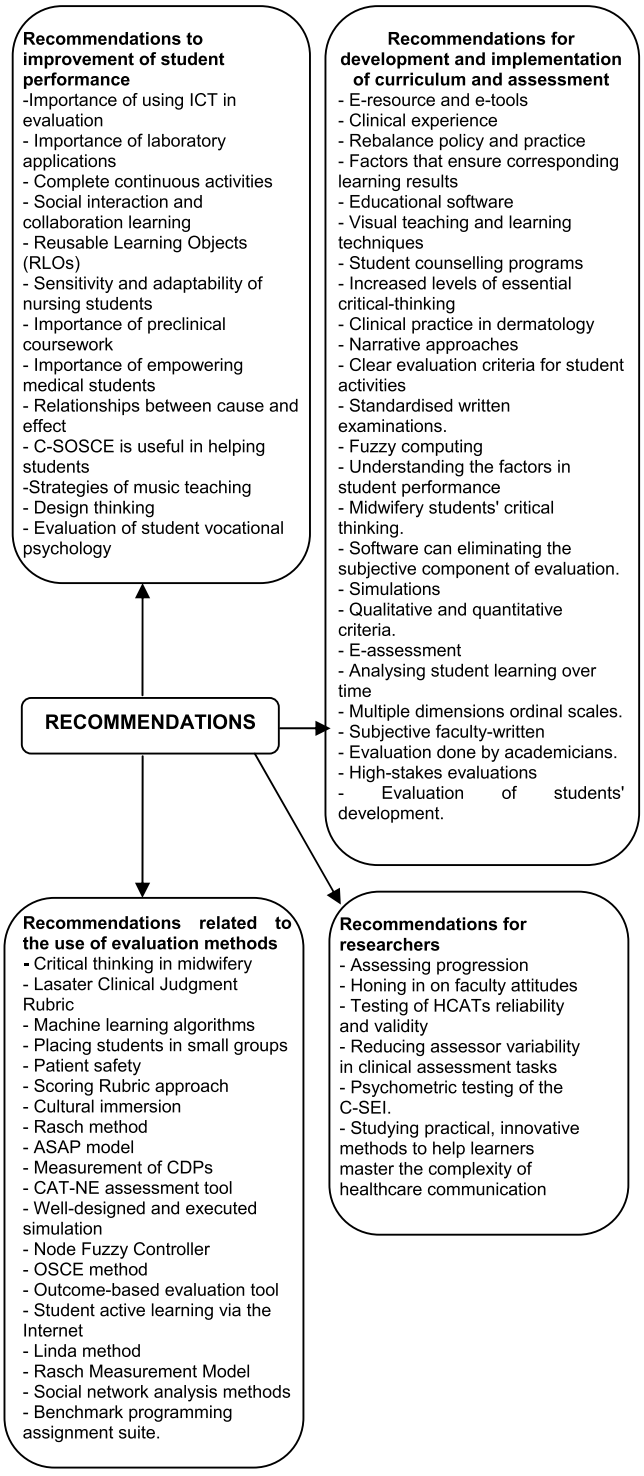

FIGURE 11. Recommendations for student evaluation.

The findings of [89] confirmed that dental schools emphasized the significance of preclinical coursework when training competent dentists. Reference [93] stated that empowering medical students to raise awareness in their communities and pushing them to serve as health advocates are important. Further, they should thoroughly understand new

medical events and gain more in-depth knowledge regarding all aspects of the event. The relationship between causes and effects in traction therapy should be considered such that each student can enhance the use of event-oriented thinking as a formative assessment and sole medical science problems [171]. C-SOSCE can help students combine the knowledge and skill before they graduate and the benefit serves evaluative strategies [77]. The improvement of music teaching strategies should be considered to support improved understanding and control of the situation by each student [151]. The evaluation of student vocational psychology can be achieved by training the students in various positions of learning to enhance their understanding of themselves in choosing suitable careers [174].

2) RECOMMENDATIONS FOR RESEARCHERS

This section includes recommendations for future research that could enhance student assessment and evaluation of overall learning and educational process. The assessment progresses in stages of team development and recognizes each activity could assist teams in accomplishing the targeted stage [8]. Reference [81] mentioned that research on faculty attitudes is needed and the evaluation of grading simulation is beyond a pass/no-pass rating scale and objective testing of student competencies in the clinical setting after participation in simulated experiences. Reference [76] confirmed that further testing HCAT reliability and validity is necessary, including inter-rater reliability, as well as measuring its usefulness in interprofessional simulations. What can be done to reduce assessor variability in clinical assessment tasks should be determined, as these methods have significant potential impact for future practices in assessment as well as on future practitioners [84]. Comparisons amongst YCTD and CT tools, such as Health Sciences Reasoning Test, and further assessments of the measurement variance with the YCTD in various Asian countries and various characteristics, are recommended [63]. Further research could considerably appreciate the internalized standard addressed by practicing professional skills and to inform the clarification of standards, development of assessment processes and instrument and design of assessor training for WBA [44]. Practical, innovative methods should be studied to help learners master the complexity of healthcare communication and develop excellent communication skills that would meet current and future competency-oriented accreditation standards [3].

3) RECOMMENDATIONS FOR THE DEVELOPMENT AND IMPLEMENTATION OF CURRICULUM AND ASSESSMENT

This section presents the recommendations in developing and implementing the curriculum and assessment and evaluation of students.

The following group of recommendations is related to curriculum development and implementation.

The students are adaptable to their learning context, and the sequence of courses is subordinate to a pedagogical style. Subsequently, a certain method increases the deep learning of students and should be combined in the development of undergraduate nursing programmes [67]. Reference [90] confirmed that e-resource is a useful supplement to lectures and interactive features. In [142], the use of each e-tool has natural interface that hides the technical complexities and mixes real and digital worlds, and thus, this technique has a positive impact on the creative process, especially in the environment design of education and industry. In [27], clinical experience has high value in learning. Policies and practices should be rebalanced by strengthening the emphasis on rights in the curriculum [168]. Reference [22] indicated that the orientation of students in reaching the target goal, consciousness and verbal intelligence are factors that emphasize the gain of correspondent learning result. Reference [100] recommended that educational software and related teaching approach should be developed to provide an overall successful learning system, which may be applied by lecturers elsewhere. Implementing new visual teaching and learning techniques that promote knowledge-based development and reinforcement of medicinal chemistry concepts with clinical relevance [120] is essential. Instructors must initiate student counseling programmes for further guidance [149]. Reference [15] confirmed that nurses need to increase each level of essential critical thinking. Student performance in objective structured clinical examination enhances medical school curriculum needs to increase the importance of clinical practice in dermatology [42]. Narrative approaches could enhance career guidance [161].

Other works provided recommendations on evaluation. Reference [167] confirmed that clear evaluation criteria should be available for student activities. Reference [30] encouraged educators to place attention on the use of tools, especially those who heavily depend on standardized written examination. [171] mentioned that fuzzy computing is effective in student performance evaluation. Reference [28] recommended that understanding each factor in student performance can improve surgical clerkship experiences. Preceptor perception of midwifery student CT in practice should be measured [14] and the development of CT of undergraduate midwifery students is important [14]. The use of software could eliminate the subjective components of evaluations, which is an additional dense coverage of the knowledge taught with questions, and reduce the time of grading and evaluation of students by examiners [118]. The design and use of simulations grounded in learning theory and an assessment model have good effects [81]. The qualitative and quantitative criteria in student evaluation submitted by teachers could be simply magnified (linguistic variables and fuzzy rules) depending on high level of adaptability [12]. Successful implementation of e-assessment at a university to the benefit of students, academic staff and institution requires commitment from all partners [39]. Student learning should be analyzed over time by using a practical and valid approach [82]. Any subjective faculty-written commentary in student assessment should be considered [32]. An evaluation done by academicians other than student supervisors in the

industry is vital to assess student performance in criteria [12]. The adoption of high-stake evaluations has a profound effect on nursing education [73]. The evaluation of student development has a key part in higher education evaluation and is an important criterion to evaluate its quality [1].

VI. LIMITATIONS

Although the database sources used in this review are reliable and cover a wide group, identification was difficult. In addition, a limitation on the timeliness of the review is caused by the increasing progress in this area. Moreover, the studies at a specific period in this vital field do not necessarily reflect actual usage or impact. Instead, the data merely show the response of the research community to the field, which is the objective of this review.

VII. CONCLUSIONS

A recent approach in student evaluation is the wave of evaluation criteria adopted by different domains and fields according to a convenient study case. Research in this trend is already active although its shape and outlines are still not understood and insights into what is actually going on in this emerging line are needed in the current stage. This article aimed to contribute such insights through surveying and taxonomizing the literature. Specific patterns could be drawn from the mass of works on student evaluation. Roughly, the research can be classified into two distinct categories of reviews and research on evaluation criteria and domains. Special areas received increased attention from researchers (e.g. medical domain) as well as few criteria (e.g. GPA and skills). This finding typically reflects the types of available evaluation criteria but clearly indicates gaps in domains and criteria. Researchers have expressed the student evaluation concerns in the literature and many have suggested recommendations to resolve existing and anticipated challenges. Such list could open many opportunities for research in this trend. People would continue to adopt new criteria and domains as they appear. Therefore, for researchers to keep abreast of this race, they should look at the next thing. The next thing in student evaluation taxonomy based on methods is to classify studies according to the approaches adopted by researchers into various connected domains (e.g. biotechnology).

REFERENCES

[1] L. N. Zhang and B. Yang, "Research on the evaluation of students' development based on support vector machine," *Appl. Mech. Mater.*, vols. 651-653, pp. 2502–2505, Sep. 2014.

[2] S. R. Hamidi, Z. A. Shaffiei, S. M. Sarif, and N. Ashar, "Exploratory study of assessment in teaching and learning," in *Proc. Int. Conf. Res. Innov. Inf. Syst. (ICRIIS)*, Nov. 2013, pp. 398–403.

[3] F. W. Kron *et al.*, "Using a computer simulation for teaching communication skills: A blinded multisite mixed methods randomized controlled trial," *Patient Educ. Counseling*, vol. 100, no. 4, pp. 748–759, 2017.

[4] M. C. San and F. Rogan, "Assessing students' English language proficiency during clinical placement: A qualitative evaluation of a language framework," *Nurse Educ. Today*, vol. 35, no. 6, pp. 771–776, 2015.

[5] M. C. Tews, R. W. Treat, and M. Nanes, "Increasing completion rate of an M4 emergency medicine student end-of-shift evaluation using a mobile electronic platform and real-time completion," *Western J. Emergency Med.*, vol. 17, no. 4, pp. 478–483, 2016.

[6] S. Walia and D. Marks-Maran, "Leadership development through action learning sets: An evaluation study," *Nurse Educ. Pract.*, vol. 14, no. 6, pp. 612–619, 2014.

[7] S. Chandrasekaran, J. M. Long, and M. A. Joordens, "Evaluation of student learning outcomes in fourth year engineering mechatronics through design based learning curriculum," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, Oct. 2015, pp. 1–7.

[8] M. Z. Farland, P. B. Barlow, T. L. Lancaster, and A. S. Franks, "Comparison of answer-until-correct and full-credit assessments in a team-based learning course," *Amer. J. Pharmaceutical Educ.*, vol. 79, no. 2, 2015, Art. no. 21.

[9] R. O'Brocta and S. Swigart, "Student perceptions of a top 200 medication course utilizing active learning techniques," *Currents Pharmacy Teach. Learn.*, vol. 5, no. 1, pp. 49–53, 2013.

[10] P. Molins-Ruano, P. Rodriguez, S. Atrio, and G. M. Sacha, "Modelling experts' behavior with e-valUAM to measure computer science skills," *Comput. Hum. Behav.*, vol. 61, pp. 378–385, Aug. 2016.

[11] C. Bennett, "Assessment rubrics: Thinking inside the boxes," *Learn. Teaching*, vol. 9, no. 1, pp. 50–72, 2016.

[12] N. Yusof and S. N. F. MohdFauzi, "Students' performance in practical training: Academicians evaluation," *Procedia-Social Behav. Sci.*, vol. 93, pp. 1275–1280, Oct. 2013.

[13] H. Ahn and H.-Y. Kim, "Implementation and outcome evaluation of high-fidelity simulation scenarios to integrate cognitive and psychomotor skills for Korean nursing students," *Nurse Educ. Today*, vol. 35, no. 5, pp. 706–711, 2015.

[14] A. G. Carter *et al.*, "Development and psychometric testing of the Carter assessment of critical thinking in midwifery (preceptor/mentor version)," *Midwifery*, vol. 34, pp. 141–149, Mar. 2016.

[15] S. A. Paul, "Assessment of critical thinking: A Delphi study," *Nurse Educ. Today*, vol. 34, no. 11, pp. 1357–1360, 2014.

[16] J. D. Kibble, "Best practices in summative assessment," in *Proc. Amer. Physiol. Soc.*, 2017, pp. 110–119.

[17] K. M. Hiller and A. Waterbrook, and K. Waters, "Timing of emergency medicine student evaluation does not affect scoring," *J. Emergency Med.*, vol. 50, no. 2, pp. 302–307, 2016.

[18] K. L. Linaker, "Student evaluations, outcomes, and national licensure examinations in radiology education: A narrative review of the literature," *J. Chiropractic Humanities*, vol. 22, no. 1, pp. 17–21, 2015.

[19] A. Rauf, M. S. Shamim, S. M. Aly, T. Chundrigar, and S. N. Alam, "Formative assessment in undergraduate medical education: Concept, implementation and hurdles," *J. Pak. Med. Assoc.*, vol. 64, no. 64, pp. 72–75, 2014.

[20] K. Helminen, K. Coco, M. Johnson, H. Turunen, and K. Tossavainen, "Summative assessment of clinical practice of student nurses: A review of the literature," *Int. J. Nursing Stud.*, vol. 53, pp. 308–319, Jan. 2016.

[21] K. A. Rasmani, N. A. Shahari, J. M. Garibaldi, and Q. Shen, "Practicality issues in using fuzzy approaches for aggregating students' academic performance," *Procedia-Social Behav. Sci.*, vol. 83, pp. 398–402, Jul. 2013.

[22] R. Lile and C. Bran, "The assessment of learning outcomes," *Procedia-Social Behav. Sci.*, vol. 163, pp. 125–131, Dec. 2014.

[23] S. Paturusi, Y. Chisaki, and T. Usagawa, "Development and evaluation of online quizzes to enhance learning performance: A survey of student assessment through MOODLE in Indonesian National University," in *Proc. Int. Conf. Inf., Commun. Technol. Syst. (ICTS)*, Sep. 2014, pp. 211–216.

[24] M. Casey, D. Wilkinson, J. Fitzgerald, D. Eley, and J. Connor, "Clinical communication skills learning outcomes among first year medical students are consistent irrespective of participation in an interview for admission to medical school," *Med. Teacher*, vol. 36, no. 7, pp. 640–642, 2014.

[25] A. E. Hill, B. J. Davidson, S. McAllister, J. Wright, and D. G. Theodoros, "Assessment of student competency in a simulated speech-language pathology clinical placement," *Int. J. Speech-Lang. Pathol.*, vol. 16, no. 5, pp. 464–475, 2014.

[26] J. Smith *et al.*, "Using cultural immersion as the platform for teaching Aboriginal and Torres Strait Islander health in an undergraduate medical curriculum," *Rural Remote Health*, vol. 15, no. 3, pp. 1–9, 2015.

[27] J. A. Myers *et al.*, "NBME subject examination in surgery scores correlate with surgery clerkship clinical experience," *J. Surgical Educ.*, vol. 71, no. 2, pp. 205–210, 2014.
[28] C. M. Reid, D. Y. Kim, J. Mandel, A. Smith, and V. Bansal, "Correlating surgical clerkship evaluations with performance on the National Board of Medical Examiners examination," *J. Surgical Res.*, vol. 190, no. 1, pp. 29–35, 2014.
[29] C. Roberts, N. Shadbolt, T. Clark, and P. Simpson, "The reliability and validity of a portfolio designed as a programmatic assessment of performance in an integrated clinical placement," *BMC Med. Educ.*, vol. 14, no. 1, p. 197, 2014.
[30] N. M. Dubosh, J. Fisher, J. Lewis, and E. A. Ullman, "Faculty evaluations correlate poorly with medical student examination performance in a fourth-year emergency medicine clerkship," *J. Emergency Med.*, vol. 52, no. 6, pp. 850–855, Jun. 2017.
[31] K. Wirth, B. Malone, C. Turner, R. Schulze, W. Widmann, and A. Sanni, "A structured teaching curriculum for medical students improves their performance on the national board of medical examiners shelf examination in surgery," *Amer. J. Surgery*, vol. 209, no. 4, pp. 765–770, 2015.
[32] K. L. Butler *et al.*, "Surgery clerkship evaluations are insufficient for clinical skills appraisal: The value of a medical student surgical objective structured clinical examination," *J. Surgical Educ.*, vol. 74, no. 2, pp. 286–294, 2017.
[33] P. A. Sutton, J. Mason, D. Vimalachandran, and S. McNally, "Attitudes, motivators, and barriers to a career in surgery: A national study of UK undergraduate medical students," *J. Surgical Educ.*, vol. 71, no. 5, pp. 662–667, 2014.
[34] F. Behnaz, G. Mohammadzade, R. S. Mousavi-e-Roknabadi, and M. Mohammadzadeh, "Assessment of knowledge, attitudes and practices regarding tuberculosis among final year students in Yazd, central Iran," *J. Epidemiology Global Health*, vol. 4, no. 2, pp. 81–85, 2014.
[35] C. Carrion, M. Soler, and M. Aymerich, "Professionalism evaluation in medical students," *Procedia-Social Behav. Sci.*, vol. 116, pp. 1880–1884, Feb. 2014.
[36] C. Carrion and M. M. S. Aymerich, "Association of professional training evaluation to content evaluation," *Procedia-Social Behav. Sci.*, vol. 93, pp. 1911–1915, 2013.
[37] M. R. Sajid, A. F. Laheji, F. Abothenain, Y. Salam, D. AlJayar, and A. Obeidat, "Can blended learning and the flipped classroom improve student learning and satisfaction in Saudi Arabia?" *Int. J. Med. Educ.*, vol. 7, pp. 281,–285, Sep. 2016.
[38] M. Blazek, J. Bess, L. Hirshbein, C. Chiang, D. Sastry, and D. Ravindranath "The short-answer vignette examination (SAVE): An assessment tool for the core psychiatry clerkship," *Acad. Psychiatry*, vol. 38, no. 5, pp. 615–618, 2014.
[39] U. G. Singh and J. M. Wassermann, "A story of a journey in implementing an E-assessment system at a South African University," *Afr. Educ. Rev.*, vol. 13, nos. 3–4, pp. 1–16, 2016.
[40] S. J. Durning *et al.*, "Are commonly used premedical school or medical school measures associated with board certification?" *Mil. Med.*, vol. 180, no. 4S, pp. 18–23, 2015.
[41] S.-W. Liao *et al.*, "Comparison of proficiency in an anesthesiology course across distinct medical student cohorts: Psychometric approaches to test equating," *J. Chin. Med. Assoc.*, vol. 77, no. 3, pp. 150–154, 2014.
[42] D. Saceda-Corralo *et al.*, "Objective structured clinical examination as an assessment tool for clinical skills in dermatology," *Actas Dermo-Sifiliográficas*, vol. 108, no. 3, pp. 237–243, 2017.
[43] E. J. N. Stupple, F. A. Maratos, J. Elander, J. Elander, K. Y. F. Cheung, and A. V. Aubeeluck, "Development of the critical thinking toolkit (CriTT): A measure of student attitudes and beliefs about critical thinking," *Thinking Skills Creativity*, vol. 23, pp. 91–100, Mar. 2017.
[44] C. Poole and J. Boland, "What influences assessors' internalised standards?" *Radiography*, vol. 22, no. 2, pp. e99–e105, May 2016.
[45] S. Hasan, R. AlSabbagh, R. AlHumaidi, M. AlMallah, and F. Khan, "Objective structured clinical examination (OSCE) in assessing pharmacy students' competence of asthma management in English and Arabic languages," *Pharmacy Educ.*, vol. 16, Oct. 2016.
[46] A. C. Hirsch and H. S. Parihar, "A capstone course with a comprehensive and integrated review of the pharmacy curriculum and student assessment as a preparation for advanced pharmacy practice experiences," *Amer. J. Pharmaceutical Educ.*, vol. 78, no. 10, 2014, Art. no. 192.
[47] J. W. Skelley, "A small group assignment gives students a novel opportunity to demonstrate current clinical controversies in a self-care course," *Amer. J. Pharmaceutical Educ.*, vol. 78, no. 10, p. 193, 2014.
[48] L. Baumann-Birkbeck *et al.*, "Benefits of e-learning in chemotherapy pharmacology education," *Currents Pharmacy Teach. Learn.*, vol. 7, no. 1, pp. 106–111, 2015.
[49] M. Hernick, A. C. Hylton, and M. Justice, "A process for curricular improvement based on evaluation of student performance on milestone examinations," *Amer. J. Pharmaceutical Educ.*, vol. 80, no. 9, p. 159, 2016.
[50] S. K. Peasah and L. L. Marshall, "The use of debates as an active learning tool in a college of pharmacy healthcare delivery course," *Currents Pharmacy Teach. Learn.*, vol. 9, no. 3, pp. 433–440, 2017.
[51] M. DeGeeter, S. R. Taylor, J. A. Wilson, P. T. Rodgers, and K. I. Leadon, "Student pharmacist perceptions of therapeutic lifestyle change (TLC) counseling abilities and personal health status," *Currents Pharmacy Teach. Learn.*, vol. 8, no. 6, pp. 750–756, 2016.
[52] Y.-C. Chen, M. E. Kiersma, and A. Abdelmageed, "Evaluation of student perceptions of standardized patient simulation on patient counseling confidence during introductory pharmacy practice experiences," *Currents Pharmacy Teach. Learn.*, vol. 7, no. 6, pp. 811–818, 2015.
[53] M. Z. Farland, A. S. Franks, P. B. Barlow, A. S. Rowe, and M. Chisholm-Burns, "Assessment of student learning patterns, performance, and long-term knowledge retention following use of didactic lecture compared to team-based learning," *Currents Pharmacy Teach. Learn.*, vol. 7, no. 3, pp. 317–323, 2015.
[54] M. Lizer, G. Ayers, R. J. Pierce, and T. Elsobky, "Pharmacy student preferences on instructional strategies in a mental health elective," *Currents Pharmacy Teach. Learn.*, vol. 8, no. 1, pp. 133–138, 2016.
[55] M. Zomorodi *et al.*, "Impact of an interprofessional population health course and clinical immersion experience: Students and practice outcomes," *J. Interprofessional Educ. Pract.*, vol. 9, pp. 91–94, Dec. 2017.
[56] K. L. DiDonato, V. L. Ruehter, L. E. Odum, A. M. Brownfield, and C. C. Lindsey, "Assessing student confidence and understanding with the immunization process through implementation of an immunization protocol at a school of pharmacy," *Currents Pharmacy Teach. Learn.*, vol. 5, no. 5, pp. 417–423, 2013.
[57] S. M. Galal *et al.*, "Assessing student knowledge, confidence, accuracy, and proficiency in providing medicare part D assistance," *Currents Pharmacy Teach. Learn.*, vol. 9, no. 2, pp. 272–281, 2017.
[58] M. R. Hanselin, C. W. Martell, S. Knutsen, R. C. Haight, and R. Moote, "Development of a capstone course to improve student confidence and pharmacotherapy knowledge prior to advanced pharmacy practice experiences," *Currents Pharmacy Teach. Learn.*, vol. 8, no. 3, pp. 323–331, 2016.
[59] C. Redmond *et al.*, "Using reusable learning objects (RLOs) in wound care education: Undergraduate student nurse's evaluation of their learning gain," *Nurse Educ. Today*, vol. 60, pp. 3–10, Jan. 2018.
[60] S.-J. Kim, S. Kim, K.-A. Kang, J. Oh, and M.-N. Lee, "Development of a simulation evaluation tool for assessing nursing students' clinical judgment in caring for children with dehydration," *Nurse Educ. Today*, vol. 37, pp. 45–52, Feb. 2016.
[61] S.-Y. Chao, H.-Y. Liu, M.-C. Wu, M. J. Clark, and J.-Y. Tan, "Identifying critical thinking indicators and critical thinker attributes in nursing practice," *J. Nursing Res.*, vol. 21, no. 3, pp. 204–210, 2013.
[62] J. Naber and T. H. Wyatt, "The effect of reflective writing interventions on the critical thinking skills and dispositions of baccalaureate nursing students," *Nurse Educ. Today*, vol. 34, no. 1, pp. 67–72, 2014.
[63] H. Shin, C. G. Park, and H. Kim, "Validation of Yoon's critical thinking disposition instrument," *Asian Nursing Res.*, vol. 9, no. 4, pp. 342–348, 2015.
[64] A. M. Elliott, "Identifying professional values in nursing: An integrative review," *Teach. Learn. Nursing*, vol. 12, no. 3, pp. 201–206, 2017.
[65] M. E. Cockerham, "Effect of faculty training on improving the consistency of student assessment and debriefing in clinical simulation," *Clin. Simul. Nursing*, vol. 11, no. 1, pp. 64–71, 2015.
[66] V. R. Terry, C. Moloney, L. Bowtell, and P. C. Terry, "Online intravenous pump emulator: As effective as face-to-face simulation for training nursing students," *Nurse Educ. Today*, vol. 40, pp. 198–203, May 2016.
[67] K. Falk, H. Falk, and E. J. Ung, "When practice precedes theory—A mixed methods evaluation of students' learning experiences in an undergraduate study program in nursing," *Nurse Educ. Pract.*, vol. 16, no. 1, pp. 14–19, 2016.
[68] H. Skúladóttir and M. H. Svavarsdóttir, "Development and validation of a clinical assessment tool for nursing education (CAT-NE)," *Nurse Educ. Pract.*, vol. 20, pp. 31–38, Sep. 2016.

[69] P. Rakbamrung, J. Puekkong, and P. Thepnuan, "Students' understanding physics concept of traction therapy," *Procedia-Social Behav. Sci.*, vol. 197, pp. 135–139, Jul. 2015.

[70] L. A. Myler and K. Seurynck, "Student evaluation of simulation in a new hospital-based simulation center," *Nursing Educ. Perspect.*, vol. 37, no. 6, pp. 335–336, 2016.

[71] M. F. Zasadny and R. M. Bull, "Assessing competence in undergraduate nursing students: The amalgamated students assessment in practice model," *Nurse Educ. Pract.*, vol. 15, no. 2, pp. 126–133, 2015.

[72] D. Markenson, S. Woolf, I. Redlener, and M. Reilly, "Disaster medicine and public health preparedness of health professions students: A multidisciplinary assessment of knowledge, confidence, and attitudes," *Disaster Med. Public Health Preparedness*, vol. 7, no. 5, pp. 499–506, 2013.

[73] M. H. Oermann, S. Kardong-Edgren, and M. A. Rizzolo, "Towards an evidence-based methodology for high-stakes evaluation of nursing students' clinical performance using simulation," *Teaching Learn. Nursing*, vol. 11, no. 4, pp. 133–137, 2016.

[74] S. Smith, S. Farra, R. T. Eyck, and M. Bashaw, "Development of an instrument to measure nursing student teamwork skills," *Clin. Simul. Nursing*, vol. 11, no. 12, pp. 507–512, 2015.

[75] F.-Y. Lin, W.-W. Wu, H.-R. Lin, and T.-Y. Lee, "The learning experiences of student nurses in pediatric medication management: A qualitative study," *Nurse Educ. Today*, vol. 34, no. 5, pp. 744–748, 2014.

[76] S. H. Campbell, M. P. Pagano, E. R. O'Shea, C. Connery, and C. Caron, "Development of the health communication assessment tool: Enhancing relationships, empowerment, and power-sharing skills," *Clin. Simul. Nursing*, vol. 9, no. 11, pp. e543–e550, 2013.

[77] E.-H. Ha, "Undergraduate nursing students' subjective attitudes to curriculum for simulation-based objective structured clinical examination," *Nurse Educ. Today*, vol. 36, pp. 11–17, Jan. 2016.

[78] R. Potenza *et al.*, "Effectiveness of an education program on donation and transplant aimed at students of the nursing degree course," *Transplantation Proc.*, vol. 47, no. 7, pp. 2097–2101, 2015.

[79] A. W. Mikasa, T. F. Cicero, and K. A. Adamson, "Outcome-based evaluation tool to evaluate student performance in high-fidelity simulation," *Clin. Simul. Nursing*, vol. 9, no. 9, pp. e361–e367, 2013.

[80] R. H. Najjar, A. Docherty, and N. Miehl, "Psychometric properties of an objective structured clinical assessment tool," *Clin. Simul. Nursing*, vol. 12, no. 3, pp. 88–95, 2016.

[81] L. Veltri, J. R. Kaakinen, C. Shillam, E. Arwood, and K. Bell, "Controlled postpartum–newborn simulation with objective evaluation exchanged for clinical learning," *Clin. Simul. Nursing*, vol. 12, no. 5, pp. 177–186, 2016.

[82] L. S. Ball and L. Kilger, "Analyzing nursing student learning over time in simulation," *Nursing Educ. Perspect.*, vol. 37, no. 6, pp. 328–330, 2016.

[83] C. A. Clark, "Evaluating nurse practitioner students through objective structured clinical examination," *Nursing Educ. Perspect.*, vol. 37, no. 6, pp. 328–330, 2016.

[84] M. Daly, Y. Salamonson, P. J. Glew, and B. Everett, "Hawks and doves: The influence of nurse assessor stringency and leniency on pass grades in clinical skills assessments," *Collegian*, vol. 24, no. 5, pp. 449–454, Oct. 2017.

[85] N. M. George, D. M. Drahnak, D. L. Schroeder, and E. D. Katrancha, "Enhancing prelicensure nursing students' use of an electronic health record," *Clin. Simul. Nursing*, vol. 12, no. 5, pp. 152–158, 2016.

[86] M. Salzmann-Erikson, M. Bjuhr, and G. Mårtensson, "Developing, implementing, and evaluating the educational module students active learning via Internet observations (SALIO) in undergraduate nursing education," *Perspect. Psychiatric Care*, vol. 53, no. 2, pp. 104–110, 2017.

[87] J. F. Giddens, L. Lauzon-Clabo, P. G. Morton, P. Jeffries, B. McQuade-Jones, and S. Ryan, "Re-envisioning clinical education for nurse practitioner programs: Themes from a national leaders' dialogue," *J. Prof. Nursing*, vol. 30, no. 3, pp. 273–278, 2014.

[88] M. P. Judge, E. C. Polifroni, A. T. Maruca, M. E. Hobson, A. Leschak, and H. Zakewicz, "Evaluation of students' receptiveness and response to an interprofessional learning activity across health care disciplines: An approach toward team development in healthcare," *Int. J. Nursing Sci.*, vol. 2, no. 1, pp. 93–98, 2015.

[89] B. C. Velayo, P. C. Stark, S. E. Eisen, and G. Kugel "Using dental students' preclinical performance as an indicator of clinical success," *J. Dental Educ.*, vol. 78, no. 6, pp. 823–828, 2014.

[90] S. Mehta, F. Clarke, and P. S. Fleming, "An assessment of student experiences and learning based on a novel undergraduate e-learning resource," *Brit. Dental J.*, vol. 221, no. 3, pp. 131–136, 2016.

[91] M. Navazesh, S. K. Rich, N. B. Chopiuk, and R. G. Keim "Triple jump examinations for dental student assessment," *J. Dental Educ.*, vol. 77, no. 10, pp. 1315–1320, 2013.

[92] K. Wilbur, N. Hassaballa, O. S. Mahmood, and E. K. Black, "Describing student performance: A comparison among clinical preceptors across cultural contexts," *Med. Educ.*, vol. 51, no. 4, pp. 411–422, 2017.

[93] A. Al-Mohrej and S. Agha, "Are Saudi medical students aware of middle east respiratory syndrome coronavirus during an outbreak?" *J. Infection Public Health*, vol. 10, no. 4, pp. 388–395, 2017.

[94] S. J. Sample *et al.*, "Students' perception of case-based continuous assessment and multiple-choice assessment in a small animal surgery course for veterinary medical students," *Veterinary Surgery*, vol. 43, no. 4, pp. 388–399, 2014.

[95] M. A. Hassanien, "A six step approach for developing computer based assessment in medical education," *Med. Teacher*, vol. 35, pp. S15–S19, Apr. 2013.

[96] B. Vaughan and T. Morrison, "Assessment in the final year clinical practicum of an Australian osteopathy program," *Int. J. Osteopathic Med.*, vol. 18, no. 4, pp. 278–286, 2015.

[97] B. Vaughan, "Confirmatory factor analysis of the study process questionnaire in an Australian osteopathy student population," *Int. J. Osteopathic Med.*, vol. 20, pp. 62–67, Jun. 2016.

[98] D. W. L. Ho, T. L. Whitehill, and V. Ciocca, "Performance of speech-language pathology students in problem-based learning tutorials and in clinical practice," *Clin. Linguistics Phonetics*, vol. 28, nos. 1–2, pp. 102–116, 2014.

[99] B. Pejcinovic, D. Duncan, P. K. Wong, M. Faust, and G. Recktenwald, "Assessment of student preparedness for freshman engineering courses through assessment of math background," in *Proc. Frontiers Educ. Conf. (FIE)*, Oct. 2014, pp. 1–5.

[100] E. Johnsen, M. Nilsen, E. Hjelseth, and C. Merschbrock, "Exploring a simple visualization tool for improving conceptual understanding of classical beam theory," *Procedia Eng.*, vol. 164, pp. 172–179, 2016.

[101] J. Hambach, C. Diezemann, M. Tisch, and J. Metternich, "Assessment of students' lean competencies with the help of behavior video analysis–are good students better problem solvers?" *Procedia CIRP*, vol. 55, pp. 230–235, 2016.

[102] S. Kadry, "Systematic assessment of student outcomes in mathematics for engineering students," in *Proc. IEEE Global Eng. Educ. Conf. (EDUCON)*, Mar. 2015, pp. 782–788.

[103] D. Fonseca, S. Villagrasa, F. Vails, E. Redondo, A. Climent, and L. Vicent, "Engineering teaching methods using hybrid technologies based on the motivation and assessment of student's profiles," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, Oct. 2014, pp. 1–8.

[104] S.-H. Jin, K. Song, S. H. I. N. Hyoung, and S. Shin, "A performance-based evaluation rubric for assessing and enhancing engineering design skills in introductory engineering design courses," *Int. J. Eng. Educ.*, vol. 31, no. 4, pp. 1007–1020, 2015.

[105] N. Khatimin, A. Zaharim, A. A. Aziz, J. Sahari, and R. A. O. K. Rahmat, "Setting the standard for project design course using Rasch measurement model," in *Proc. Global Eng. Educ. Conf. (EDUCON)*, Mar. 2013, pp. 1062–1065.

[106] M. Samarakou, P. Prentakis, D. Mitsoudis, D. Karolidis, and S. Athinaios, "Application of fuzzy logic for the assessment of engineering students," in *Proc. IEEE Global Eng. Educ. Conf. (EDUCON)*, Apr. 2017, pp. 646–650.

[107] A. I. Mecwan, D. B. Shah, and B. D. Fataniya, "Innovations in evaluation: An integral part of outcome based education," in *Proc. 5th Nirma Univ. Int. Conf. Eng. (NUiCONE)*, Nov. 2015, pp. 1–5.

[108] H. Jackson, K. Tarhini, S. Zelmanowitz, and C. Hatfield, "Developing a simple and effective method of assessing civil engineering student outcomes," in *Proc. Frontiers Educ. Conf. (FIE)*, Oct. 2012, pp. 1–6.

[109] M. M. U. Faiz, U. B. Mansoor, S. M. Asad, and K. Mahmood, "Using faculty course assessment report for the assessment of an associate degree course in engineering technology program," in *Proc. IEEE 6th Conf. Eng. Educ. (ICEED)*, Dec. 2014, pp. 73–78.

[110] Y. Luo and W. Wu, "Sustainable design with BIM facilitation in project-based learning," *Procedia Eng.*, vol. 118, pp. 819–826, 2015.

[111] I. Damaj and J. Yousafzai, "Simple and accurate student outcomes assessment: A unified approach using senior computer engineering design experiences," in *Proc. IEEE Global Eng. Educ. Conf. (EDUCON)*, Apr. 2016, pp. 204–211.

[112] Z. Yıldız and A. F. Baba, "Evaluation of student performance in laboratory applications using fuzzy decision support system model," in *Proc. Global Eng. Educ. Conf. (EDUCON)*, Apr. 2014, pp. 1023–1027.
[113] O. Macek and M. Kom'rek "Evaluation of student teamwork," in *Proc. IEEE 25th Conf. Softw. Eng. Educ. Training (CSEE&T)*, Apr. 2012, pp. 130–133.
[114] Y. Kalaani and R. J. Haddad, "A unified approach to the assessment of student learning outcomes in electrical engineering programs," Tech. Rep., 2014.
[115] S. Rana and R. Garg, "Application of hierarchical clustering algorithm to evaluate students performance of an institute," in *Proc. 2nd Int. Conf. Comput. Intell. Commun. Technol. (CICT)*, Feb. 2016, pp. 692–697.
[116] K. Mahmood, K. M. Khan, K. S. Khan, and S. Kiani, "Implementation of outcome based education in Pakistan: A step towards Washington Accord," in *Proc. IEEE 7th Int. Conf. Eng. Educ. (ICEED)*, Nov. 2015, pp. 166–170.
[117] N. Rajaee, M. F. M. Sabri, K. H. Ping, and M. H. Husin, "Assessment of student learning outcomes in digital signal processing course in PKEK, UNIMAS," in *Proc. IEEE 5th Conf. Eng. Educ. (ICEED)*, Dec. 2013, pp. 21–25.
[118] G. Vuc, F. Baloi, and M. Litcanu, "Adapting methods of student evaluation and grading in electrical power engineering," *Procedia-Social Behav. Sci.*, vol. 191, pp. 147–151, Jun. 2015.
[119] H.-J. Chen, J.-L. She, C.-C. Chou, Y.-M. Tsai, and M.-H. Chiu, "Development and application of a scoring rubric for evaluating students' experimental skills in organic chemistry: An instructional guide for teaching assistants," *J. Chem. Educ.*, vol. 90, no. 10, pp. 1296–1302, 2013.
[120] K. A. El Sayed and C. T. Chelette, "Laboratory exercises to teach clinically relevant chemistry of antibiotics," *Amer. J. Pharmaceutical Educ.*, vol. 78, no. 2, 2014, Art. no. 37.
[121] K. M. Elkins and K. L. Murphy, "Use of the online version of an ACS general chemistry exam: Evaluation of student performance and impact on the final exam," in *Proc. Technol. Assessment Strategies Improving Student Learn. Chem.*, vol. 1235, Nov. 2016, pp. 211–223.
[122] S. Zhou *et al.*, "Assessment of scientific reasoning: The effects of task context, data, and design on student reasoning in control of variables," *Thinking Skills Creativity*, vol. 19, pp. 175–187, Mar. 2016.
[123] F. M. Ibrahim, A. A. Shariff, and R. M. Tahir, "Using Rasch model to analyze the ability of pre-University students in vector," *AIP Conf. Proc.*, vol. 1682, p. 030009, Oct. 2015.
[124] F. M. Ibrahim, Y. Z. Zubairi, A. A. Aziz, and A. Zaharim, "Using the Rasch model to assess examination beyond students' scores," in *Proc. IEEE Global Eng. Educ. Conf. (EDUCON)*, Apr. 2012, pp. 1–5.
[125] N. G. Holmes, J. Ives, and D. A. Bonn, "The impact of targeting scientific reasoning on student attitudes about experimental physics," in *Proc. Phys. Educ. Res. Conf.*, 2014, pp. 119–122.
[126] L. S. Kiani and C. Menke, "Refining scientific writing skills with feedback that works for students and instructors," in *Proc. Educ. Training Opt. Photon. Opt. Soc. Amer.*, 2015, Paper MEE03.
[127] T. L. Larkin, "Breaking with tradition: Using the conference paper as a case for alternative assessment in physics," in *Proc. Int. Conf. Interact. Collaborative Learn. (ICL)*, Sep. 2013, pp. 744–751.
[128] S. Banerjee, C. Ramanathan, and N. J. Rao, "An approach to automatic evaluation of higher cognitive levels assessment items," in *Proc. IEEE 3rd Int. Conf. Innov. Technol. Educ. (MITE)*, Oct. 2015, pp. 342–347.
[129] S. Felemban, M. Gardner, and V. Callaghan, "An event detection approach for identifying learning evidence in collaborative virtual environments," in *Proc. 8th Comput. Sci. Electron. Eng. (CEEC)*, Sep. 2016, pp. 42–47.
[130] B. Kakani, D. Dalal, and A. Dabhi, "Improved solution on students answer sheet assessment using fuzzy rules," in *Proc. Conf. Adv. Signal Process. (CASP)*, Jun. 2016, pp. 435–439.
[131] D. Baneres, "Towards a particular prediction system to evaluate student's success," in *Proc. Int. Conf. P2P, Parallel, Grid, Cloud Internet Comput.* Springer, 2016, pp. 935–945.
[132] V. P. Gurupur, G. P. Jain, and R. Rudraraju, "Evaluating student learning using concept maps and Markov chains," *Expert Syst. Appl.*, vol. 42, no. 7, pp. 3306–3314, 2015.
[133] K. K. Sharma, K. Banerjee, C. Mandal, and I. Vikas, "A benchmark programming assignment suite for quantitative analysis of student performance in early programming courses," in *Proc. IEEE 3rd Int. Conf. MOOCs, Innov. Technol. Educ. (MITE)*, Oct. 2015, pp. 199–203.
[134] A. Cain, "Developing assessment criteria for portfolio assessed introductory programming," in *Proc. IEEE Int. Conf. Teach., Assessment Learn. Eng. (TALE)*, Aug. 2013, pp. 55–60.
[135] H. A. M. Abdeljaber and S. Ahmad, "Program outcomes assessment method for multi-academic accreditation bodies: Computer science program as a case study," *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 5, pp. 23–35, 2017.
[136] T. Maria, P. Dimitris, F. Garifallos, G. Athanasios, and M. Roumeliotis, "Collaboration learning as a tool supporting value co-creation. Evaluating students learning through concept maps," *Procedia-Social Behav. Sci.*, vol. 182, pp. 375–380, May 2015.
[137] A. S. García, M. T. García-Álvarez, and B. Moreno, "Analysis of assessment opportunities of learning spaces: On-line versus face to face methodologies," *Comput. Hum. Behav.*, vol. 30, pp. 372–377, Jan. 2014.
[138] I. Simonova and P. Poulova, "Assessment preferences of IT and management students," in *Proc. IEEE Conf. e-Learn., e-Manage. e-Services (IC)*, Oct. 2016, pp. 58–63.
[139] A. González-Marcos, F. Alba-Elías, F. Navaridas-Nalda, and J. Ordieres-Meré, "Student evaluation of a virtual experience for project management learning: An empirical study for learning improvement," *Comput. Educ.*, vol. 102, pp. 172–187, Nov. 2016.
[140] P. He, "Evaluating students online discussion performance by using social network analysis," *Proc. 9th Int. Conf. Inf. Technol., New Gener. (ITNG)*, Apr. 2012, pp. 854–855.
[141] S. A. Abdullah, A. Yaakub, and Z. Wahil, "Evaluating students' need in using computer aided software in landscape design course," *Procedia-Social Behav. Sci.*, vol. 195, pp. 828–836, Jul. 2015.
[142] G. Salvador, M. Bañó, M. Contero, and J. Camba, "Evaluation of a distributed collaborative workspace as a creativity tool in the context of design education," in *Proc. Frontiers Educ. Conf. (FIE)*, Oct. 2014, pp. 1–7.
[143] H. Jeong, "Narrative and expository genre effects on students, raters, and performance criteria," *Assessing Writing*, vol. 31, pp. 113–125, Jan. 2017.
[144] P. Crosthwaite, S. Boynton, and S. Cole, "Exploring rater conceptions of academic stance and engagement during group tutorial discussion assessment," *J. English Acad. Purposes*, vol. 28, pp. 1–13, Jul. 2017.
[145] B. Whitlock and N. Ebrahimi, "Beyond the library: Using multiple, mixed measures simultaneously in a college-wide assessment of information literacy," *College Res. Libraries*, vol. 77, no. 2, pp. 236–262, 2016.
[146] M.-H. Wang, C.-S. Wang, C.-S. Lee, S.-W. Lin, and P.-H. Hung, "Type-2 fuzzy set construction and application for adaptive student assessment system," in*Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2014, pp. 888–894.
[147] A. R. Daif and M. A. Rizkaa, "An enhanced model for monitoring learners' performance in a collaborative e-Learning environment," in *Proc. 2nd Int. Conf. e-Learn. e-Technol. Educ. (ICEEE)*, Sep. 2013, pp. 313–317.
[148] G. Stoet and D. C. Geary, "Students in countries with higher levels of religiosity perform lower in science and mathematics," *Intelligence*, vol. 62, pp. 71–78, May 2017.
[149] Y.-H. Hu, C.-L. Lo, and S.-P. Shih, "Developing early warning systems to predict students' online learning performance," *Comput. Hum. Behav.*, vol. 36, pp. 469–478, Jul. 2014.
[150] Y. Chen, C.-C. Pan, G.-K. Yang, and J. Bai, "Intelligent decision system for accessing academic performance of candidates for early admission to University," in *Proc. 10th Int. Conf. Natural Comput. (ICNC)*, Aug. 2014, pp. 687–692.
[151] P. Blanco-Piñeiro, M. P. Díaz-Pereira, and A. Martínez, "Common postural defects among music students," *J. Bodywork Movement Therapies*, vol. 19, no. 3, pp. 565–572, 2015.
[152] M. Misut and M. Misutova, "Software solution improving productivity and quality for big volume students' group assessment process," *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 4, pp. 175–190, 2017.
[153] M. Lotriet, H. C. Erasmus, and S. N. Mostert, "Tutoring targets–the challenge of evaluating success," *J. New Gener. Sci.*, vol. 14, no. 1, pp. 63–76, 2016.
[154] S. H. J. Petrudi, M. Pirouz, and B. Pirouz, "Application of fuzzy logic for performance evaluation of academic students," in *Proc. 13th Iranian Conf. Fuzzy Syst. (IFSC)*, Aug. 2013, pp. 1–5.
[155] H. Heflin, J. Shewmaker, and J. Nguyen, "Impact of mobile technology on student attitudes, engagement, and learning," *Comput. Educ.*, vol. 107, pp. 91–99, Apr. 2017.

[156] M. Worrachananun, F. Yan, and B. Zhaoming, "A qualitative assessment of student-centered approach: High-level question assessment method among chinese and macanese students in Macao," in *Proc. Int. Conf. Educ. Innov. Through Technol. (EITT)*, Oct. 2015, pp. 160–164.
[157] D. Tafiadis *et al.*, "Comparison of voice handicap index scores between female students of speech therapy and other health professions," *J. Voice*, vol. 31, no. 5, pp. 583–588, 2017.
[158] C. Sin and M. Manatos, "Student assessment in Portugal: Academic practice and Bologna policy," *Higher Educ. Policy*, vol. 27, no. 3, pp. 323–340, 2014.
[159] E. Viennet and L. Petrucci, "Monitoring students performances in French Institutes of Technology using the ScoDoc software," in *Proc. Inf. Technol. Based Higher Educ. Training (ITHET)*, Sep. 2014, pp. 1–6.
[160] M. I. M. Saad, S. Baharom, and S. E. Mokhsein, "Scientific reasoning skills based on socio-scientific issues in the biology subject," *Int. J. Adv. Appl. Sci.*, vol. 4, no. 3, pp. 13–18, 2017.
[161] R. Lengelle, F. Meijers, R. Poell, and M. Post, "Career writing: Creative, expressive and reflective approaches to narrative identity formation in students in higher education," *J. Vocational Behav.*, vol. 85, no. 1, pp. 75–84, 2014.
[162] J. Kohoutek, "European standards for quality assurance and institutional practices of student assessment in the UK, the Netherlands and the Czech Republic," *Assessment Eval. Higher Edu.*, vol. 39, no. 3, pp. 310–325, 2014.
[163] T. Sundeen, K. M. V. Garland, and W. D. Wienke, "A multi-year evaluation of student perceptions of University and special education Doctoral Websites," *Teacher Educ. Special Educ.*, vol. 39, no. 4, pp. 259–275, 2016.
[164] R. E. W. Thomas, T. L. Teel, and B. L. Bruyere, "Seeking excellence for the land of paradise: Integrating cultural information into an environmental education program in a rural Hawai'ian community," *Stud. Educ. Eval.*, vol. 41, pp. 58–67, Jun. 2014.
[165] S. N. Ingoley and J. W. Bakal, "A novel method for inferring strict and lenient marks into normal marks using fuzzy logic," in *Proc. Nirma Univ. Int. Conf. Eng. (NUiCONE)*, Nov. 2013, pp. 1–6.
[166] M. Nallasamy, "A study on threshold concepts in teaching and learning in TAFE-industry training," in *Proc. Int. Conf. Teach., Assessment Learn. Eng. (TALE)*, Dec. 2014, pp. 419–422.
[167] M. Ito, N. Naoe, A. Imazawa, and O. Matsushita, "Introduction of adapting design thinking into the education in Kanazawa Technical College," in *Proc. IEEE 7th Int. Conf. Eng. Educ. (ICEED)*, Nov. 2015, pp. 25–28.
[168] B. Macfarlane, "The performative turn in the assessment of student learning: A rights perspective," *Teach. Higher Edu.*, vol. 21, no. 7, pp. 839–853, 2016.
[169] Meenakshi and P. Nagar, "Fuzzy logic based expert system for students' performance evaluation," in *Proc. 2nd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, Mar. 2015, pp. 803–808.
[170] T. Mahboob, S. Irfan, and A. Karamat, "A machine learning approach for student assessment in E-learning using Quinlan's C4.5, Naive Bayes and Random Forest algorithms," in *Proc. 19th Int. Multi-Topic Conf. (INMIC)*, Dec. 2016, pp. 1–8.
[171] A. Barlybayev, A. Sharipbay, G. Ulyukova, T. Sabyrov, and B. Kuzenbayev, "Student's performance evaluation by fuzzy logic," *Procedia Comput. Sci.*, vol. 102, pp. 98–105, 2016.
[172] S. N. Ingoley and J. W. Bakal, "Evaluating students' performance using four-node fuzzy controller," in *Proc. Nirma Univ. Int. Conf. Eng. (NUiCONE)*, Nov. 2013, pp. 1–6.
[173] B. Chakraborty and M. Sinha, "Student evaluation model using Bayesian network in an intelligent e-learning system," *J. Inst. Integrative Omics Appl. Biotechnol.*, vol. 7, no. 2, pp. 51–60, 2016.
[174] Y. Liu, C. Han, L. Huang, B. Wang, and Z. Zhu, "Research and development of student assessment system based on knowledge, ability and mentality," in *Proc. 7th Int. Conf. Comput. Sci. Educ. (ICCSE)*, Jul. 2012, pp. 1829–1832.
[175] J. Addo-Atuah, A. Dutta, and C. Kovera, "A global health elective course in a PharmD curriculum," *Amer. J. Pharmaceutical Educ.*, vol. 78, no. 10, p. 187, 2014.
[176] B. Dance *et al.*, "An information literacy snapshot: Authentic assessment across the curriculum," *College Res. Libraries*, vol. 76, no. 2, pp. 170–187, 2015.
[177] K. Murray, K. McKenzie, and M. Kelleher, "The evaluation of a framework for measuring the non-technical ward round skills of final year nursing students: An observational study," *Nurse Educ. Today*, vol. 45, pp. 87–90, Oct. 2016.
[178] L. Ezechil, "New perspectives on evaluation and certifying the competences of higher education graduates," *Procedia-Social Behav. Sci.*, vol. 76, pp. 7–12, Apr. 2013.

OMAR ZUGHOUL received the B.S degree in computer information system from Yarmouk university, Irbid, Jordan, in 2011, and the M.Sc. degree in software engineering from UNITEN University, Malaysia, in 2014. He is currently pursuing the Ph.D. degree with Universiti Pendidikan Sultan Idris, Tanjung Malim, Malaysia. His research areas are student evaluation, student assessment, student appraisal, and educational domains.

FAIEZ MOMANI received the B.Sc. degree in computer information system from Philadelphia University, Amman, Jordan, in 2011, and the M.Sc. degree in information technology from Universiti Tenaga Nasional, Bangi, Malaysia, in 2014. He is currently pursuing the Ph.D. degree with Universiti Pendidikan Sultan Idris, Tanjung Malim, Malaysia. His research areas are student evaluation, student assessment, student appraisal, and educational domains.

O. H. ALMASRI received the B.Sc. degree in computer systems engineering from Mamoun Private University, Syria, in 2009, and the M.Sc. degree in information technology from Universiti Tenaga Nasional, Bangi, Malaysia, in 2014, where he is currently pursuing the Ph.D. degree. His research areas are student evaluation, student assessment, student appraisal, and educational domains.

A. A. ZAIDAN received the B.Eng. degree (Hons.) in computer engineering from the University of Technology, Baghdad, Iraq, in 2004, the M.Sc. degree in data communications and computer network from the University of Malaya, Malaysia, in 2009, and the Ph.D. degree in artificial intelligence from Multimedia University, Malaysia, in 2013. He was the leader or a member for many funded research projects. He is currently a Senior Lecturer with the Department of Computing, University Pendidikan Sultan Idris. He has published more than 160 papers at various international conferences and journals. His research areas are student evaluation, student assessment, student appraisal, and educational domains.

B. B. ZAIDAN received the B.Sc. degree in applied mathematics from Al-Nahrain University, Baghdad, Iraq, in 2004, and the M.Sc. degree in data communications and information security from the University of Malaya, Malaysia, in 2009. He was the leader or a member for many funded research projects. He is currently a Senior Lecturer with the Department of computing, University Pendidikan Sultan Idris. He has published more than 160 papers at various international conferences and journals. His research interests include student evaluation, student assessment, student appraisal, and educational domains.

M. A. ALSALEM received the B.Sc. and M.Sc. degrees in management information system from Mosul University, Mosul, Iraq, in 2007 and 2012, respectively. He is currently pursuing the Ph.D. degree with the Faculty of Art, Computing and Creative Industry, University Pendidikan Sultan Idris, Malaysia. His research interests include the measurement of Student evaluation, student assessment, student appraisal, and educational domains.

O. S. ALBAHRI received the B.Sc. degree in computer science from the Al Turath University College, Baghdad, Iraq, in 2011 and the M.Sc. degree in computer science and communication from Arts, Sciences and Technology University in Lebanon, Beirut, Lebanon, in 2014. He is currently pursuing the Ph.D. degree with Universiti Pendidikan Sultan Idris, Tanjung Malim, Malaysia. His research areas are student evaluation, student assessment, student appraisal, and educational domains.

A. S. ALBAHRI received the B.Sc. degree in computer science from Al-Mamon University College, Baghdad, Iraq, in 2000, and the M.Sc. degree in computer science and communication from Arts, Sciences and Technology University in Lebanon, Beirut, Lebanon. He is currently pursuing the Ph.D. degree with Universiti Pendidikan Sultan Idris, Tanjung Malim, Malaysia. He is also a Lecturer with Iraqi Commission for Computers and Informatics. His research areas are student evaluation, student assessment, student appraisal, and educational domains.

M. HASHIM received the B.Sc. and M.Sc. degrees in information technology from University Kebangsaan Malaysia, in 1998 and 1999, respectively. She is currently a Senior Lecturer with the Department of Computing, University Pendidikan Sultan Idris. Her research areas are student evaluation, student assessment, student appraisal, and educational domains.