Article

Assessing and Benchmarking Learning Outcomes of Robotics-Enabled STEM Education

S. M. Mizanoor Rahman [ID]

Department of Intelligent Systems and Robotics, University of West Florida, Pensacola, FL 32514, USA; smrahman@uwf.edu

Abstract: Experienced middle school mathematics and science teachers were recruited for a pilot study. The teachers separately responded to a survey related to determining expected learning outcomes based on their traditional teaching, classroom experiences and observations, and self-brainstorming. The teachers then received training on how to design, develop, and implement robotics-enabled lessons under a design-based research approach for experiential learning, and taught robotics-enabled lessons to a selected student population in classroom settings. The teachers then responded to the survey for the robotics-enabled teaching. For each case (traditional and robotics-enabled), the survey responses were analyzed, and a set of expected learning outcomes of math and science lessons was derived separately. The thematic analysis results showed that the expected learning outcomes for the robotics-enabled lessons were not only related to the educational gains (content knowledge) observed in traditional teaching, but also to the improvements in the behavioral, social, scientific, cognitive, and intellectual aptitudes of the students. Then, a set of metrics and methods were proposed for assessing the learning outcomes separately. To validate the assessment metrics and methods, teachers from different schools taught two selected robotics-enabled lessons (one math, one science) to same grade students, and separately assessed the learning outcomes of each student using the proposed metrics and methods. The learning outcomes were then compared and benchmarked between schools and subjects. The results of a user study with the teachers showed user acceptance, effectiveness, and suitability of the assessment metrics and methods. The proposed scheme of assessing learning outcomes can be used to assess and justify the benefits and advantages of robotics-enabled STEM education, benchmark the outcomes, help improve teaching preparations, motivate decision-makers to confer on robotics-enabled STEM education and curricula development, and promote robotics-enabled STEM education.

Keywords: robotics; K-12 STEM education; learning outcomes; assessment; benchmarking; design-based research; experiential learning; thematic analysis; classroom observations; brainstorming

1. Introduction

Students in K-12 levels need to learn STEM concepts that are fully or partly abstract in nature [1,2]. Learning abstract STEM topics in young ages may limit the learners' comprehension abilities [1–6]; especially, such abstract practices may increase the cognitive workload of the learners, and decrease their computational thinking abilities [2]. This problem can become more severe when lower grade students (e.g., K-8) attempt to learn the abstract concepts [3]. This problem can also persist with upper grade students in college levels, especially with the freshmen or with other grade college students who attempt to learn new concepts for the first time. It is assumed that suitably designed tangible learning platforms may provide active and experiential learning opportunities to students that may facilitate them having kinesthetic learning experiences and making the abstract STEM concepts easy to comprehend [2]. In this regard, it is posited that application of robotic and mechatronic devices as tangible learning platforms may be a pragmatic choice to illustrate abstract STEM concepts to students [1–7]. The reasons behind

choosing robotic devices as teaching (or pedagogical) and learning tools are that they may be cost-effective, they can show movements that help create live illustrations, and the design and development of the robotic structures are adjustable that can be used to create many illustrations easily in the shortest possible time [1–6]. The adjustability and flexibility in the design of robotic teaching platforms may help teachers adjust the robotic devices with the actual needs of individual learners, and thus may help ensure equity and maintain an inclusive learning environment [2,6]. In this way, robotics may serve as a pedagogical tool, which is more than a teaching aid. Another reason behind the robotic platforms is their potential abilities to intrinsically or extrinsically motivate the learners (students), teachers (instructors), institutions, and the parents [2]. Robotic platforms may also be helpful towards cognitive apprenticeship [6], project or problem-based learning (PBL), classroom engagement [8], inquiry-based learning (IBSME), situated leaning and situated cognition, etc. Considering the above potential advantages of robotics-enabled STEM education, efforts towards proposing various approaches of robotics-enabled STEM education are gaining prominence among researchers [1–7].

However, robotics-enabled STEM education still did not receive expected priority and attention, and the current contributions are limited and immature [1]. More innovative efforts towards robotics-enabled experiential kinesthetic learning are necessary. Especially, robotics-enabled illustrations created to illustrate abstract STEM concepts need to be more meaningful, logical, practical, relevant, and appropriate for the targeted student grades [7–13]. The illustrations should be easy to understand, implementable within the timeframe and resources in actual classroom settings, and safe to the users (learners and teachers) [6]. Moreover, illustrations should not create any misconceptions in the learners instead of easing the comprehension of the abstract concepts, and illustrations should not overrule any true preconceptions [2]. However, development of robotics-enabled STEM education addressing these requirements and expectations is still not sufficient as seen in the literature [7–13], except a very few preliminary initiatives [1–6,14–16].

An important question in implementing robotics-enabled STEM education is, "How can we justify that the robotics-enabled STEM education actually enhances the learning outcomes of the students (learners), and what are the scope and elements of the learning outcomes?" This question is very logical because teachers and education decision-makers may not find interest and rationale in implementing robotics-enabled STEM education if they are not enough convinced that robotics-enabled STEM lessons truly and positively impact the learning outcomes of the students, and the outcomes are better than that of the traditional education [5,17–26]. The students and their parents/guardians may turn their faces from robotics-enabled lessons if the students are not benefited from the applications of robotics in their lessons in terms of learning outcomes, and the outcomes are measurable [27,28]. Hence, an appropriate scheme (appropriate and comprehensive assessment methods and metrics [29–43]) is necessary to identify the scope and elements of expected learning outcomes and assess and justify the outcomes of robotics-enabled STEM education [3]. Such a scheme is also necessary to benchmark the learning outcomes among students, grades, schools, school zones, and school districts, and to show the differences of outcomes between robotics-enabled and traditional education [2,44]. However, such a comprehensive assessment scheme comprising of appropriate assessment metrics and methods to assess and benchmark the learning outcomes of robotics-enabled STEM education is not observed in the literature, except for a very few preliminary efforts [1–6].

It is true that various assessment methods and metrics of assessing learning outcomes are already available in the literature, e.g., [17–26]. However, these methods and metrics may not be suitably applicable to robotics-enabled STEM education because the robotics-enabled STEM education includes robots as a technology-based pedagogical tool that is usually not considered when the state-of-the-art learning assessment methods and metrics are proposed, such as [17–26]. It is believed that the inclusion of robots as a pedagogical tool creates different teaching and learning scenarios and contexts that may create different expectations about the learning outcomes in the learners, teachers, school administration,

parents, and communities [2,17–26,45–49]. If so, it requires an appropriately designed and customized assessment scheme for robotics-enabled STEM education [50–59]. In addition, there may have differences in the assessment schemes between student grades, subjects, etc. Furthermore, the assessment scheme needs to be validated in actual classroom settings for its practicality and generalization. However, the attempt to propose an appropriate, customized, comprehensive, and validated assessment scheme comprising of appropriate assessment methods and metrics for assessing learning outcomes of robotics-enabled STEM education is still a future work.

Hence, the objective of this article is to attempt to propose an appropriate, customized, and comprehensive assessment scheme comprising of appropriate assessment methods and metrics for assessing the learning outcomes of robotics-enabled STEM education for K-12 levels, and to validate the assessment scheme for its practicality and generalization. For the simplicity, in this article, the efforts will be limited to middle school math and science lessons. However, the results, in principle, may be applicable to robotics-enabled STEM lessons for K-12 and collegiate levels as well. The learning vision is that the derived assessment methods and metrics can guide robotics-enabled lesson designers to design, implement, and predict expected learning outcomes of the lessons, which can help them distinguish the benefits of robotics-enabled lessons from that of traditional lessons, and thus can highlight the importance of robotics-enabled lessons [60–64]. The vision is also that the overall efforts can enhance the learning outcomes significantly.

2. Analysis of Related Works

A plethora of research reports are found in the existing literature that clearly show the growing interests in research in robotics-enabled STEM education [1–7,13]. The applications of repeated evaluation and feedback approach were proposed and verified to assess and optimize design, development, and implementation performance of a professional development program for in-service middle school teachers for teaching K-12 STEM lessons using robotics-enabled illustrations [1]. The prerequisites (the expected qualifications, attitudes and aptitudes) of K-12 students interested in attending robotics-enabled STEM lessons were determined [2]. The prerequisites were meant to be the qualifications, attitudes, and aptitudes that the prospective students would need to have to obtain optimum benefits from robotics-enabled STEM lessons. For this purpose, the computational thinking ability of the students was identified as one of the key requirements that the students would need to have before they could attend robotics-enabled STEM lessons. It was found that the computational thinking ability of the students might also increase if the students could participate in robotics-enabled STEM lessons [2].

A teaching framework called the technological-pedagogical and content knowledge (TPACK) was applied to instruct robotics-enabled middle school mathematics and science lessons [3]. The variations in the applications of the TPACK, and the impact of the TPACK framework on teaching robotics-enabled STEM lessons with varying difficulty were investigated, and the outcomes of the robotics-enabled lessons over that of the traditionally taught lessons were compared. The results showed the superiority of the robotics-enabled teaching over the traditional teaching. The dynamic behaviors of the TPACK framework for teaching robotics-enabled STEM lessons in middle schools were explored [4]. The results showed significant variations in the effectiveness of TPACK with variations in subjects, grades, and teachers. The factors affecting the trust of students and teachers in robots for robotics-enabled middle school STEM lessons were determined [5]. It was found that there was significant impacts of the trust of the students in the robots on their learning outcomes in their robotics-enabled STEM lessons. The systems approach to analyzing the design-based research strategy in robotics-enabled middle school STEM lessons was proposed, and its effectiveness was justified [6]. The effectiveness of cognitive apprenticeship approach in conjunction with the systematic design-based research was confirmed.

It was found that robots were applied to enhance the learning effectiveness of English language in elementary school students [7]. As was reported, a framework utilizing LEGO

robots was developed to enhance problem solving ability in students [8]. The authors found that the robotics-enabled teaching enhanced student engagement in classrooms. The authors found that the use of LEGO robots was effective to create interests among high school students in their STEM lessons [9]. The review results showed that the social robots could be useful in education as they could be used as robot tutors or robotic peer learners [10]. It was argued that the social robots were proved effective in improving the cognitive and affective abilities of students. It was reported that the learning outcomes were similar to those of human teachers tutoring similar lessons. It might happen because of the interactive embodiment and physical presence of the social robots that the traditional non-robotic teaching and learning technologies and facilities could not provide. A review study on the applications of robotics in STEM education especially on young children education was conducted [11]. It was showed that there was a strong trend in the effectiveness of robotics-enabled education on children. A systematic survey to explore the educational potentials of robots and robotics-enabled lessons in school environment was conducted, and strong learning potential were found [12]. It was found that the creativity of students in higher education significantly increased through robotics-enabled STEM lessons [13].

Instructing a mechatronics course to undergraduate engineering students in colleges following the TPACK framework was proven efficacious [14]. An all-in-one robotic platform was used to instruct mechatronics fundamentals such as actuators and sensors to the students. It was found that instructing mechatronics concepts using the robotic platform seemed to enhance learning outcomes and learners' satisfaction. The 7E instructional model with the design-based research (DBR) method was proposed to design and instruct a mechatronics course for undergraduate engineering students [15]. Some robotics devices such as actuator and sensor systems were used as the pedagogical tools to instruct the mechatronics concepts, especially the fundamental concepts of actuation and sensing. It was found that the implementation of mechatronics lessons following the 7E instructional model along with the DBR method enhanced the teaching and learning outcomes and effectiveness. A few mechanical engineering concepts such as additive manufacturing (3D printing), pneumatics principles, fine machining (such as laser engraving), etc. were instructed through the applications of a robotic platform. It was found that the experiential kinesthetic learning through the applications of a robotic platform enhanced the teaching and learning outcomes and effectiveness significantly [16].

In all of the above examples, different aspects of robotics-enabled STEM education were addressed. Enhancing overall learning outcomes was considered as the main objective of implementing robotics-enabled STEM education [7–13,17–23]. A few examples such as [14–16] attempted to present the impacts of robotics-enabled STEM education on the learning outcomes. The literature shows that researchers are very active in proposing different approaches to assess the learning outcomes [17–26], including the SOLO model. The authors expressed the learning outcomes in terms of critical thinking [17]. The authors explained the learning outcomes from the perspective of the students or the learners [18]. The authors described the importance of assessing learning outcomes of students in higher education [19]. Other issues related to assessing the learning outcomes such as the definitions, thresholds, roles, integration, student perception and sustainability of the learning outcomes were presented in various ways [20–26]. However, no scheme seems to be holistic and comprehensive, rather each of those schemes focused on some part of the learning outcomes. So far, there is no generalization in the assessment methods and metrics, which cannot help benchmark the learning outcomes among different grade students, subjects, and schools. Most importantly, those state-of-the-art works did not consider robotics-enabled STEM education. It is assumed that a new paradigm in studying the learning outcomes is necessary for the lessons instructed using robotics as pedagogical tools. It is believed that a comprehensive assessment scheme could capture all the possible and relevant learning outcomes of robotics-enabled education, and thus could help use the integrated results for various purposes such as curriculum development, benchmarking, student and teacher awarding, education related policy planning and decision-making, etc.

However, the current initiatives in the literature do not focus on developing such a customized, comprehensive, and holistic assessment scheme for assessing and benchmarking the learning outcomes of robotics-enabled STEM education.

Based on the aforementioned literature reviews representing the state-of-the-art research and development activities in this field, it can be posited that there is a big gap in the state-of-the-art works regarding assessing the learning outcomes of robotics-enabled STEM education, especially in the K-12 classes. Thus, an appropriately designed comprehensive assessment and benchmarking scheme to assess and benchmark the learning outcomes of robotics-enabled STEM education is still a future work. This paper aims to contribute to this direction, and bridge the gaps in the state-of-the-art knowledge and practices of assessing the learning outcomes of robotics-enabled STEM education.

3. Research Questions

Considering the gaps in the state-of-the-art research related to learning outcomes of robotics-enabled STEM lessons (education) as above [17–26], answers to the following research questions will be sought in this paper:

Q 1: Are the methods and metrics used to assess the learning outcomes of robotics-enabled STEM lessons different for teaching different subjects (e.g., math and science) and different grade (e.g., 6–8) students?

Q 2: How are the methods and metrics used to assess the learning outcomes of robotics-enabled STEM lessons different from or similar to that of the traditionally taught non-robotics-based STEM lessons?

4. Materials and Resources

In total, 20 math and 20 science teachers from 20 middle schools were recruited to participate in the pilot study. As the sampling procedures, for the teachers, we contacted the selected schools and circulated a recruitment notice. On the notice, we mentioned that the teachers who had good experience in math and science teaching following traditional methods but did not have experience in robotics-enabled lessons would apply. We then conducted an interview with each teacher separately, and conveyed the information regarding the duties and responsibilities of the teachers for the study. We also considered the years of experience of each teacher in teaching math or science in middle schools. We then selected the teachers who were found to be the most promising and interested in the proposed study. For students, we randomly selected students from each class of each selected teacher to participate in the pilot study.

We took consent of the students and teachers, and preserved their consents. The study was conducted following local ethical standards and principles for human subjects, and we were aware of the privacy and security principles for human subjects (students and teachers) mentioned in the ethical standards. We then trained the teachers on how to develop and implement robotics-enabled math and science lessons. We and the trained teachers together developed 10 math and 10 science lessons using robotics as a pedagogical tool. All lessons were planned to meet the state standards for middle school science and math based on the Next Generation Science Standards (NGSS) and the Common Core State Standards for Math (CCSSM) [44]. As an example, a math lesson is described as follows. The teachers used LEGO robots to create illustrations to teach number line to middle school students of grade 6 in their math lessons as exhibited in Figure 1. In such an example, a number line was drawn on the classroom floor. The number line was divided into positive and negative digits. The space between two adjacent digits had a value of |1|. A LEGO robot vehicle was programmed to move along the number line. The touch buttons were used to give addition and subtraction commands to the robot. The robot illustrated the addition or the subtraction results through its movement along the number line. For example, if it was commanded to subtract 3 from 2 (i.e., 2–3), the robot started to move from '2' towards '0', moved for 3 spaces and stopped at '−1'. Thus, the robot wanted to

illustrate that 2–3 = −1 (see Figure 1). Lesson materials such as lesson descriptions, activity sheets, instruction procedures, etc. were developed for the lesson [1–6].


Figure 1. The robot moves along a number line according to the addition or subtraction commands.

In another example, teachers put the LEGO robot vehicles at different locations on a sliding surface so that the robots could come down from higher positions to lower positions along the sliding path, as illustrated in Figure 2a. Similarly, a robot was programmed to move a block on the floor, as illustrated in Figure 2b. Those illustrations were used to teach the students about fundamentals concepts of mass, force, friction, displacement, velocity, speed, acceleration, momentum, etc. in their science lessons. Lesson materials such as lesson descriptions, activity sheets, instruction procedures, etc. were developed for the lessons. Similar examples of robotics-enabled STEM lesson design and development can be found in [1–6].
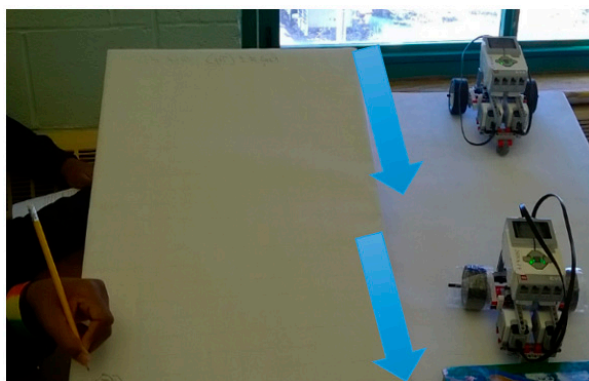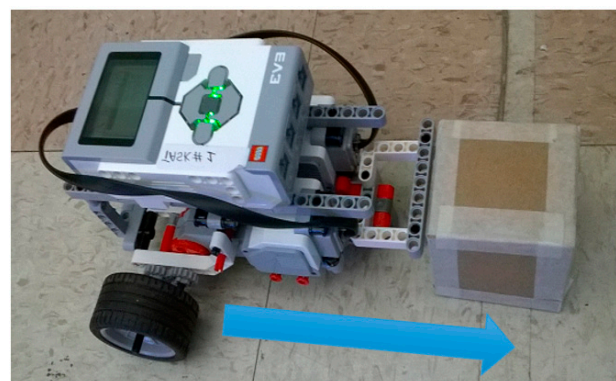
(**a**)   (**b**)

Figure 2. Robotics-enabled illustrations: (**a**) robots slide down along a surface, and (**b**) a robot moves a block on a floor to illustrate science concepts in a classroom setting.

The trained teachers separately implemented the robotics-enabled math and science lessons in actual classroom settings. Students were divided into teams, and they observed robot activities for the lessons preprogrammed by their teachers, interacted with the robots, performed lesson activities, completed activity sheets, etc. instructed by their teachers. The robotics-enabled math and science lessons, trained teachers, selected students, and the classroom settings were used as the materials and resources for the research presented herein.

5. Research Methods and Procedures

The research methods presented herein were based on surveys with the teachers [50,51], and observations on students and their classroom activities [52,62]. The research procedures included two phases (steps): (i) development of the assessment methods and metrics

for the learning outcomes of students for their robotics-enabled lessons, and (ii) validation and generalization of the assessment methods and metrics in actual classroom environment. For the first phase, a survey was conducted with the math and science teachers separately [50,51]. The survey questionnaires are given in Appendix A. The survey was conducted with each teacher twice: (i) before their trainings on robotics-enabled lessons (it was treated as the traditional or non-robotic based teaching), (ii) after their trainings on and implementation of robotics-enabled lessons (it was treated as the robotic-based/enabled teaching). The participating math and science teachers responded to the surveys separately based on their classroom experiences and observations of student activities [52,62]. They also conducted self-brainstorming to fill out the survey questionnaires [53]. The survey questionnaires were given to the responding teachers and they were allowed to take two days to think individually and respond to the survey questionnaires. Hence, the responses received from the teachers were treated as their well-thought opinions that were based on their teaching and classroom experiences. The name of each responding teacher was coded so that the true identity of the responder could not be identified while processing the responses data, as per the ethical standards. This phase of research was conducted to develop a set of assessment methods and metrics for assessing expected learning outcomes of students for their math and science lessons for both traditional and robotics-enabled scenarios.

For the second phase, 20 math teachers taught the same topic/lesson (e.g., number line) to the same grade students (e.g., grade 6 students) in their schools using robotics. Similarly, 20 science teachers taught the same topic/lesson (e.g., force/friction) to the same grade students (e.g., grade 6 students) in their schools using robotics. Then, each teacher assessed the learning outcomes of the lesson that he/she taught using the assessment methods and metrics developed in the first phase. The assessment was performed during the class, in a 1-h extra session with the participating students after the class, and during a 1-week follow up period to assess different criteria properly. The learning outcomes between schools and subjects were compared and benchmarked. Then, another user study survey was conducted with the teachers taking their opinions about the usability, practicability, and reliability of the assessment methods and metrics for assessing the learning outcomes separately [54]. The survey was based on a 7-point Likert scale where +3 was the most positive (highest) and −3 was the most negative (lowest) response [27]. The Likert scale is exhibited in Figure 3.
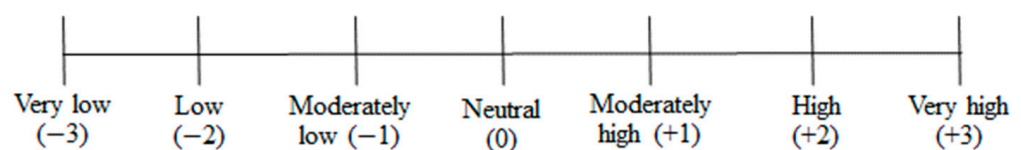
Figure 3. The Likert scale.

6. Research Results and Analyses

6.1. Determining the Assessment Criteria and Metrics

The responses of the questionnaires in Appendix A were analyzed. The responses with similar meanings for the first question for math and science lessons/teachers were tallied under different key terms separately as Tables 1 and 2 show for the math and science lessons respectively. Here, the key terms can be considered as the assessment criteria for learning outcomes, and the criteria together can be called the assessment metrics. The tables compare different criteria (key terms) proposed by the responding teachers for assessing the learning outcomes of their students for the math and science lessons between traditional and robotics-enabled teaching methods [30]. Here, the traditionally taught lessons and participants could serve as the control group for the robotics-enabled lessons group when we compared the perceived learning outcomes between the traditionally taught and the robotics-enabled lessons. The tables also show the frequencies of the responses. For example, "Problem solving ability (9)" in Table 1 for the traditional teaching

of the math lesson means that out of 20 responding math teachers, 9 teachers mentioned on their responses to the question 1 in Appendix A that the problem solving ability of the students should be considered as a criterion to assess the learning outcomes of the students for their math lessons. In other words, the problem solving ability should be an outcome of the math lesson as opined by 9 teachers out of 20.

Table 1. Comparison of the criteria proposed by the responding math teachers for assessing the learning outcomes of the students for the math lessons between traditional and robotics-enabled teaching.

Assessment Criteria of Learning Outcomes and Their Frequencies in Parentheses	
Traditional Teaching	**Robotics-Enabled Teaching**
- Test results or subject matter knowledge (20) - Engagement in the class (6) - Interpersonal (student-student-teacher) relationship (5) - Life-long learning aspiration (6) - Problem solving ability (9) - Critical thinking ability (5) - Reflexive analysis (4) - Professional ethics (2) - Decision making ability (6) - Communication skills (7) - Computational thinking ability (7)	- Test results or subject matter knowledge (20) - Engagement in the class (11) - Interpersonal (student-student-teacher) relationship (12) - Life-long learning aspiration (7) - Problem solving ability (11) - Critical thinking ability (9) - Reflexive analysis (5) - Professional ethics (3) - Decision making ability (13) - Communication skills (14) - Computational thinking ability (9) - Leadership ability (9) - Intrinsic and extrinsic motivation (12) - Trust in learning devices or robotics (4) - Attendance and time management in the class (8) - Engineering and ICT skills (7) - Hands-on and practical ability (5) - Lab skills and experiment ability (4) - Formulation of research strategy (8) - Teamwork ability (10) - Cognitive workload in learning (14) - Organizational and planning ability (7) - Troubleshooting and contingency management (6) - Creativity and innovation (9) - Interdisciplinary and multidisciplinary abilities (11) - Adapting to new situations and changes (5) - Respect for diversity and multiculturality (4) - Entrepreneurial ability (3) - Social responsibility (2)

Tables 1 and 2 show that, in general, the responding teachers expected better learning outcomes in terms of assessment criteria and their frequencies for the robotics-enabled teaching over the traditional teaching for both math and science lessons. It is assumed that the teachers' expectations of learning outcomes for the robotics-enabled teaching increased because they realized higher monetary investment, high-tech kinesthetic teaching and learning artifacts, intense classroom activities and better pedagogical clarity and transparency associated with the robotics-enabled teaching over the traditional teaching [1–6]. For example, intense classroom activities centering round the robots occurred in the classrooms for robotics-enabled teaching of math and science lessons. The students and teachers together needed to implement the lessons using robotics in the classroom environment, and the students needed to manage and complete such activities in teams within a specified timeframe. As a result, teachers might have expectations that the performed activities would create higher abilities and skills in the students related to content knowledge, leadership, social responsibility, time management, punctuality, teamwork, decision making, interpersonal relationship, classroom engagement, problem solving, critical thinking, professional ethics, communications, basic engineering, ICT, practical work, experimentation, research formu-

lation, organization and planning, troubleshooting, contingency management, adaptation to changes, creativity, innovation, etc. The well-developed robotics-based learning systems and devices might create entrepreneurial thinking in the students. The robotic device as an experiential learning tool might be itself a source of intrinsic and extrinsic motivation to the students that might engage students with their lessons, stimulate continuous and life-long learning, build trust in the learning devices, etc. The tangible and visible robotics learning tools might reduce cognitive workload of students while learning because such tools might reduce the mental demand, temporal demand, frustration, and efforts, and increase the learning performance simultaneously [28]. Students needed to utilize different multidisciplinary and interdisciplinary concepts to work with robotics-enabled lessons, and to complete the lesson activities. Students from different culture and races needed to work in teams to learn their lessons using robotics as kinesthetic learning tools. All those might enhance their interdisciplinary and multidisciplinary skills, and inculcate an inclusive, diverse, and multicultural mentality in the students.

Table 2. Comparison of the criteria proposed by the responding science teachers for assessing the learning outcomes of the students for the science lessons between traditional and robotics-enabled teaching.

Assessment Criteria of Learning Outcomes and Their Frequencies in Parentheses	
Traditional Teaching	**Robotics-Enabled Teaching**
- Test results or subject matter knowledge (20) - Engagement in the class (7) - Interpersonal (student-student-teacher) relationship (6) - Life-long learning aspiration (8) - Problem solving ability (8) - Critical thinking ability (5) - Reflexive analysis (3) - Professional ethics (3) - Decision making ability (8) - Communication skills (5) - Imagination ability (6)	- Test results or subject matter knowledge (20) - Engagement in the class (14) - Interpersonal (student-student-teacher) relationship (13) - Life-long learning aspiration (10) - Problem solving ability (15) - Critical thinking ability (12) - Reflexive analysis (6) - Professional ethics (4) - Decision making ability (15) - Communication skills (16) - Imagination ability (13) - Leadership ability (11) - Intrinsic and extrinsic motivation (16) - Trust in learning devices or robotics (5) - Attendance and time management in the class (9) - Engineering and ICT skills (9) - Hands-on and practical ability (8) - Lab skills and experiment ability (7) - Formulation of research strategy (11) - Teamwork ability (12) - Cognitive workload in learning (9) - Organizational and planning ability (9) - Troubleshooting and contingency management (7) - Creativity and innovation (11) - Interdisciplinary and multidisciplinary abilities (12) - Adapting to new situations and changes (5) - Respect for diversity and multiculturality (5) - Entrepreneurial ability (4) - Social responsibility (4)

The robotics-enabled teaching was usually more student-centered while the traditional teaching was more teacher-centered [45]. As the results in Tables 1 and 2 show, it was assumed that the paradigm shifted in the centering of the classroom activities (from teacher-centered to student-centered), which might create higher expectations of the teachers about the learning outcomes of their students for the student-centered robotics-enabled teaching [45]. The results might also indirectly reveal that the robotics-enabled teaching

should produce better learning outcomes of students in order to be admired by their teachers, parents, school administration, and school districts.

The results in Tables 1 and 2 show slight differences in the expected learning outcomes between math and science lessons for both traditional and robotics-enabled teaching methods. For example, the teachers expected computational thinking ability as the learning outcome for the math lessons. However, computational thinking ability was not expected for the science lessons. Instead, the imagination ability of students was expected for the science lessons. The reasons may be that computational thinking is more related to math than to science. However, students should have more imagination ability to imagine science concepts to comprehend them through developing or using tangible learning tools such as the robotic devices. It is further observed that the frequencies of responses from the teachers for different learning outcomes for the science lessons were greater than that for the math lessons. The reasons may be that the kinesthetic learning using robotics was expected to influence the science learning more intensely than the math learning because the math is more abstract than the science as best as it was understood while observing classroom activities associated with the lessons.

For the second question in Appendix A, 18 teachers out 20 for the math lessons opined that they did not expect different learning outcomes for different grades of middle school students. However, 19 teachers out of 20 for the science lessons did not expect different learning outcomes for different grades of middle school students. It might happen because the syllabi, standards, and depth of the education for different grades of students in middle schools are not enough different for the teachers to perceive different learning outcomes for different grade students. Hence, it is posited that the same or similar assessment criteria of learning outcomes may be used for different grades of students in middle schools. However, the differences may be easily understandable if the expected learning outcomes between middle school grades (e.g., grade 6) and high school grades (e.g., grade 10) are compared.

Then, the responses (the proposed criteria of learning outcomes) in Tables 1 and 2 for the robotics-enabled teaching were grouped under different themes separately through thematic analysis [2,55]. Table 3 shows the themes of learning outcomes for the robotics-enabled math lessons (based on Table 1) as an example. Then, the frequencies for all the criteria of each theme were added separately, as Table 4 shows. Figure 4 shows the relative and timely, create a life-long learning aspiration in the students based on their long term relationship with the tangible interactive robotic platform, and finally enhance their teamwork ability through the activities they perform in teams during the lessons centering round the robotic platform. The results show that improvements in scientific/technical, managerial/leadership, intellectual, cognitive, and social abilities of the students are also the expected outcomes of learning math lessons through a robotics-enabled teaching method. Similar results were obtained for robotics-enabled science lessons. The results in general mean that the learning outcomes of robotics-enabled math and science lessons can be treated as satisfactory if the assessment results for the mentioned criteria of learning outcomes (Tables 1 and 2) are satisfactory, and/or the assessment results for each of the outcome themes (Table 3) are satisfactory. Contribution of each theme in the total contribution. Results in Figure 4 show that the improvement in the behavioral characteristics of the students through their robotics-enabled math lessons is the most expected learning outcome. Based on the results, it is realized that the robot is not simply a pedagogical tool that can help learn the subject matter (or content knowledge), which is called here the educational outcome. Instead, the robot should generate intrinsic and extrinsic motivation in the students, enhance their trust in the robot as a learning tool, improve their physical and mental engagement with the learning platform such as the robotic platform, motivate them to attend the school regularly

Table 3. Determination of different themes of criteria of assessing the learning outcomes of the students for robotics-enabled math lessons.

Assessment Criteria (Expected Learning Outcomes)	Themes
Test results (20)	Educational
Life-long learning aspiration (7)	Behavioral
Intrinsic and extrinsic motivation (12)	Behavioral
Trust in robotics (4)	Behavioral
Engagement in class (11)	Behavioral
Class attendance and punctuality (8)	Behavioral
Adapting to new situations and changes (5)	Behavioral
Respect for diversity and multiculturality (4)	Behavioral
Professional ethics (3)	Behavioral
Teamwork ability (10)	Behavioral
Engineering and ICT skills (7)	Scientific/technical
Formulation of research strategy (8)	Scientific/technical
Hands-on and practical ability (5)	Scientific/technical
Lab skills and experiment ability (4)	Scientific/technical
Troubleshooting and contingency (6)	Scientific/technical
Interdisciplinary/multidisciplinary abilities (11)	Scientific/technical
Problem solving ability (11)	Intellectual
Reflexive analysis (5)	Intellectual
Critical thinking ability (9)	Intellectual
Computational thinking ability (9)	Intellectual
Decision making ability (13)	Intellectual
Creativity and innovation (9)	Intellectual
Cognitive workload (14)	Cognitive
Entrepreneurial ability (3)	Managerial/leadership
Communication skills (14)	Managerial/leadership
Leadership ability (9)	Managerial/leadership
Organizational and planning ability (7)	Managerial/leadership
Social responsibility (2)	Social
Interpersonal relationship (12)	Social

Table 4. Different themes of learning outcomes and the corresponding total frequencies for robotics-enabled math lessons.

Themes (of Learning Outcomes)	Total Frequencies
Educational	20
Behavioral	64
Scientific/technical	41
Intellectual	56
Cognitive	14
Managerial/leadership	33
Social	14
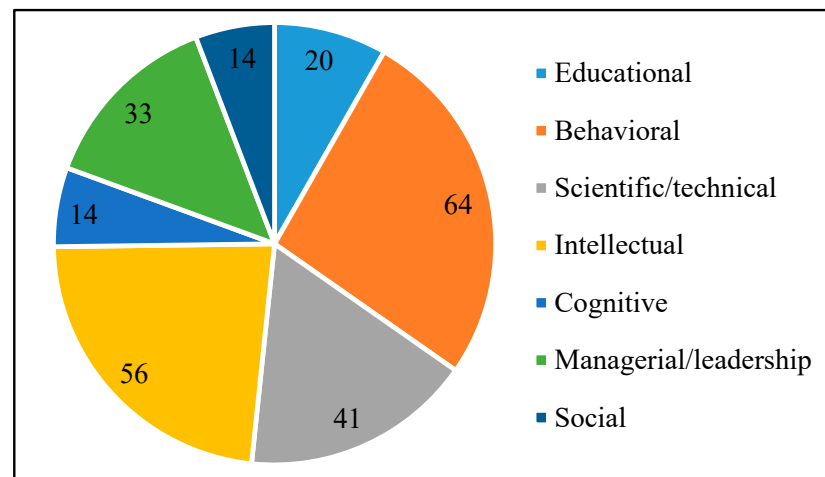
Figure 4. Relative contributions (in term of total frequencies) of different themes of expected learning outcomes of the students for learning robot-enabled math lessons.

Now, the question is what metrics are to be used to assess the mentioned criteria of learning outcomes, and how. The answer to this question is as follows. The nature of each assessment criterion in Table 1 for the robotics-enabled math lessons was critically analyzed with respect to the scenarios where the students performed the robotics-enabled activities. Then, the assessment metric for each criterion was proposed, being inspired by the existing body of knowledge on each criterion found in the literature, considering the nature of each criterion with respect to the activity scenario, and conducting brainstorming with the concerned teachers and education experts. The results are given in Table A1 (in Appendix B). Similar results were found for the robotics-enabled science lessons.

The results in Table A1 show that in some cases, the exact assessment metrics were not proposed. Those were kept open for mainly two reasons: (i) there might have multiple options for the metrics to assess those criteria of learning outcomes depending on situations and scenarios, and (ii) it was difficult to decide the metrics unless the actual scenario was known in general. In such cases, the assessment metrics would need to be determined by teachers and/or education researchers implementing robotics-enabled lessons based on their experiences, knowledge, understanding, and observations. On the other hand, the assessment methods may also be influenced by the assessment metrics, and vice versa. For example, a Likert scale is to be used as an assessment metric for a learning outcome if the outcome needs to be assessed subjectively and quantitatively [27], and vice versa. For the criteria where tests/quizzes and surveys were proposed as the assessment metrics, special quizzes/tests and surveys might need to be designed and administered. The NASA TLX and work sampling should follow the standard NASA TLX and work sampling implementation methods and materials, respectively [27,28]. For the criteria of qualitative observations, the teachers and/or education researchers will need to observe the classroom scenarios and activities, assess the learning outcomes qualitatively, and prepare a qualitative report on the assessment of each specific assessment criterion of learning outcome. Note that in actual implementation scenarios, the learning outcomes may not be favorable for all assessment criteria, which may open a road to improvements.

6.2. Validation of the Learning Outcome Assessment Methods and Metrics

For the second phase of the research, Tables A2 and A3 (in Appendix B) compare the learning outcomes assessed using different criteria and metrics proposed earlier (see Table A1) between robotics-enabled math and science lessons for different participating schools. The results show that the proposed assessment criteria and metrics of learning outcomes (Table A1) can be implemented successfully to understand the status of learning outcomes of robotics-enabled math and science lessons. The results also show that the robotics-enabled science lessons produced slightly better outcomes. However, based on

the *t*-test results between the subjects (math and science), it was found that the differences were not statistically significant ($p > 0.05$) for each criterion of learning outcome between math and science lessons. It might happen due to the reasons that the science concepts might be less abstract but more related to real-life scenarios and thus the tangible robotic platform as a learning tool might impact the science learning outcomes more intensely than the math learning outcomes. Figure 5 further exhibits the slight differences in the learning outcomes for different assessment criteria between math and science lessons.

(**a**)

(**b**)
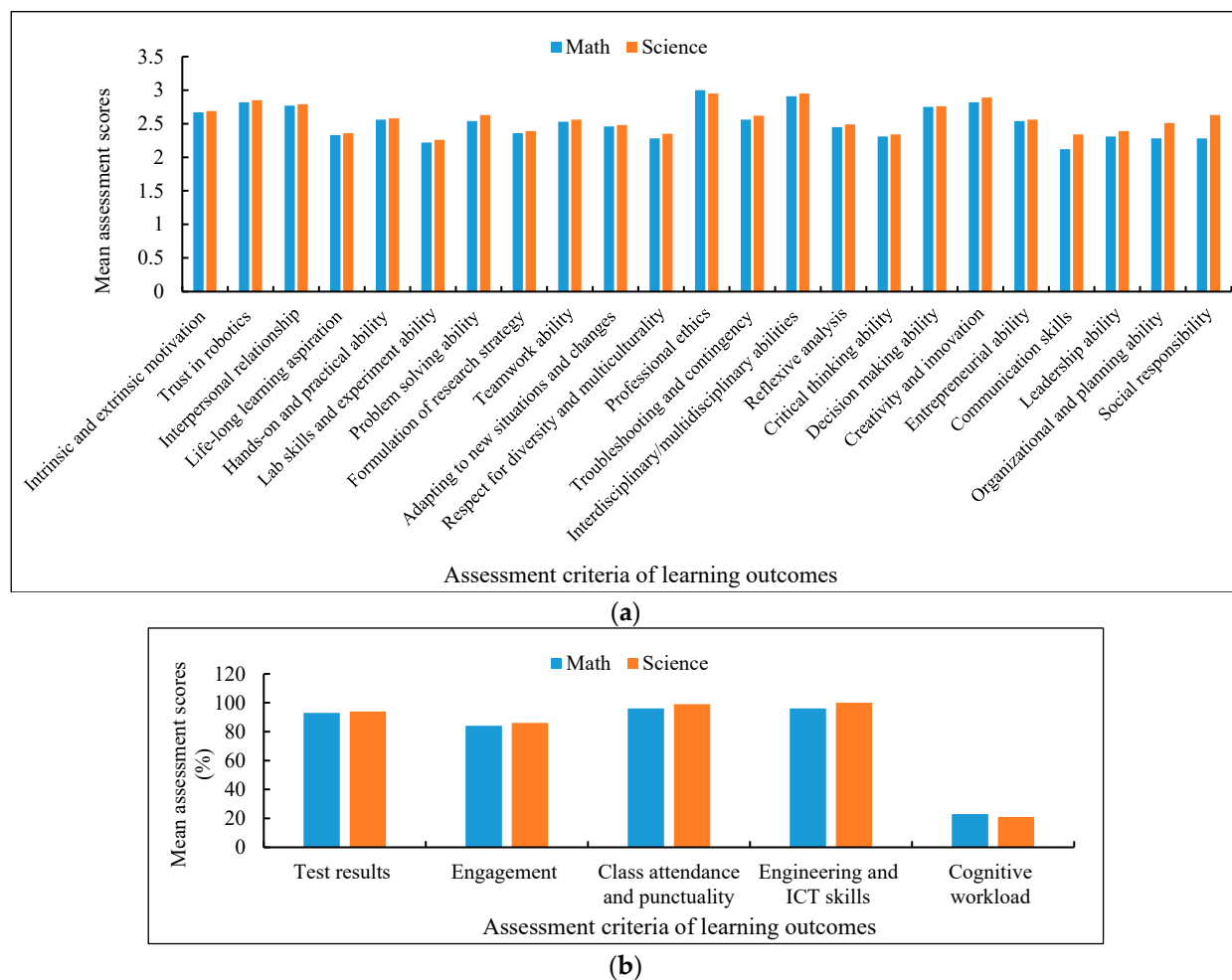
Figure 5. The mean (*n* = 20) assessment scores: (**a**) assessed based on the Likert scale (max. score +3), and (**b**) assessed as percentages of total obtainable scores (max. score 100%) for different assessment criteria of learning outcomes between the math and science lessons for school#1 as an example.

Figure 5a compares the mean assessment scores of learning outcomes assessed based on the Likert scale (max. score +3), and Figure 5b compares the mean assessment scores of learning outcomes assessed as the percentages of the total obtainable scores (max. score 100%) for different assessment criteria between the math and science lessons for school#1 as an example. These results as a whole validate the effectiveness and prove the practicality of the proposed assessment criteria and metrics of learning outcomes for the robotics-enabled math and science lessons. Therefore, the metrics can be used to compare and benchmark the learning outcomes between students, student grades, subjects, schools, and school districts.

Based on the user study (teachers' opinions) results, Figure 6 compares the usability, practicability, and reliability of the assessment criteria and metrics of learning outcomes between robotics-enabled math and science lessons. The results in Figure 6 show that the assessment scheme was proven usable, practical, and reliable as it was opined by the

users (teachers). The scheme was proven slightly better in terms of usability, practicability, and reliability for the science lesson in comparison with the math lesson. The reasons may be similar as explained earlier. These results validate the effectiveness and prove the practicality of the proposed assessment criteria and metrics for assessing the learning outcomes of robotics-enabled math and science lessons.
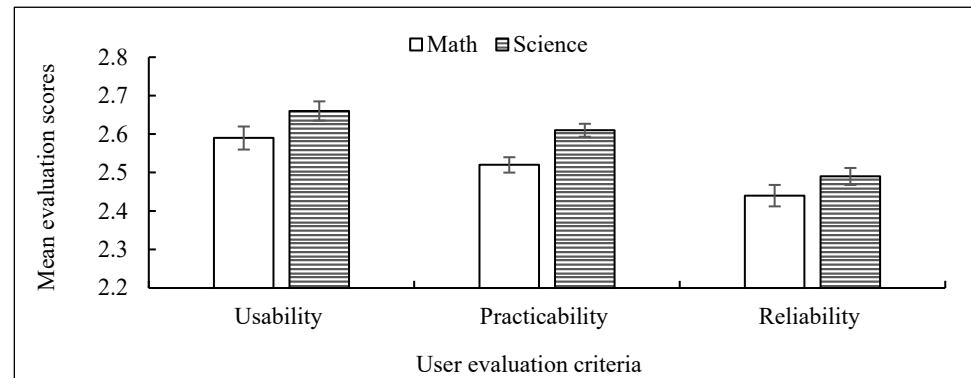
Figure 6. The mean (*n* = 20) evaluation scores for usability, practicability, and reliability of the assessment criteria and metrics of learning outcomes between robotics-enabled math and science lessons.

7. Discussion

The results of the presented study are limited in many ways. A few of the limitations can be summarized as follows: (i) the results are limited to 6–8 grade students in middle schools only, and the results may not be readily applicable to the elementary and high school grades and college levels, (ii) the study was conducted using LEGO (Mindstorms) robots, and it is yet to investigate the effectiveness of the results if other robotic platforms are used, (iii) the study was conducted with a limited number of lesson scenarios, and the results may be changed or may need to be adjusted if more lessons with different scenarios or the same lessons with different and multiple scenarios are implemented, (iv) the study was conducted with a limited number of teachers and students, and the results may need to be adjusted if greater number of teachers and students are recruited, (v) teaching experiences of teachers and previous experiences of teachers and students with robotics may also impact the results that were not considered in the presented study, (vi) the study considered only a few representative lessons from math and science, but lessons from engineering and technology need to be considered to have a clear picture about the expected learning outcomes of robotics-enabled lessons, etc. However, it is possible to address all of these limitations properly. Despite having limitations, this study can convey the preliminary information about assessing and benchmarking the expected learning outcomes of robotics-enabled STEM lessons, which is significant. The results are in line with what were found in the state-of-the-art traditional teaching and learning methods [17–26,30,46]. However, the results obtained herein augment the scope of the state-of-the-art initiatives, and increase the effectiveness of the existing results to make them suitable for teaching and learning robotics-enabled lessons.

In the integrative model of interdisciplinary learning, knowledge, modes of inquiry, and pedagogies from multiple disciplines (multiple disciplines may mean multiple majors, subjects, topics, ideas, solutions, concepts, etc.) can be brought together within the context of a single course or program or practice [56]. Students learning in this model are able to apply the knowledge gained in one discipline or subject area to different other disciplines or subject areas or concepts to deepen overall learning experiences [56]. On the other hand, active learning method asks learners to fully participate in their learning by thinking, discussing, investigating, and creating. In active learning, students/learners are asked to practice skills, solve problems, struggle with complex questions, propose solutions, and explain ideas in their own words through speaking, writing, and discussing [57]. Research shows that active learning methods are more effective than traditional lecturing for student

learning [57]. Experiential learning is another form of education closely related to active learning where students can learn though experiences [58]. It may be hypothesized that experiential learning and active learning are complementary with each other; they can be integrated and implemented with interdisciplinary learning concepts, and such an integration may be more effective and impactful than individual active learning, experiential learning, or interdisciplinary learning. Robotics can be used as a pedagogical and learning tool that can integrate and foster active learning, experiential learning, and interdisciplinary learning [56–58]. However, effective applications of such an integrated model to highly impact the STEM education are usually not observed in the literature, and the expected learning outcomes of such an integration are yet to be known. The results presented herein may inspire this multimodal integrative model of education.

8. Conclusions and Future Works

Based on a survey conducted with 40 middle school math and science teachers having experiences of developing and implementing robotics-enabled lessons, a set of expected learning outcomes of robotics-enabled STEM education (here, only the math and science education) was derived, and the metrics and methods to evaluate each outcome were proposed. The survey results showed that the expected learning outcomes were not only related to the educational gains (content knowledge), but also to the improvements in the behavioral, social, scientific, cognitive and intellectual attitudes, and aptitudes of the students. The results showed clear differences in the expected learning outcomes between the traditional and robotics-enabled experiential methods of teaching. The reasons might be the higher level investment of cognitive resources and artifacts in robotics-enabled lessons. However, the expected learning outcomes between the math and the science lessons were not so significant. The results (the set of learning outcomes, assessment metrics and methods) were then validated through actual classroom applications, and the effectiveness of the assessment methods and metrics were evaluated based on a user study with the participating teachers. The user study results proved the effectiveness of the proposed methods and metrics of learning outcomes of the robotics-enabled lessons. The main contribution of this article is the determination of the assessment and benchmarking criteria, metrics, and methods for assessing learning outcomes of robotics-enabled STEM lessons, which is novel, practical, and useful to advance robotics-enabled kinesthetic K-12 STEM education in particular, and the college-level STEM education in general. The results uphold the significance of active learning and experiential learning. The proposed evaluation scheme of learning outcomes can be used to justify the benefits and advantages of robotics-enabled STEM education, benchmark the outcomes, help improve preparations of instructors and teaching institutions and develop more effective robotic systems and demonstrations under design-based research, motivate education decision-makers to confer on robotics-enabled STEM education and curricula development, and thus can promote robotics-enabled K-16 STEM education practices. All these can help meet the learning vision of enhancing the learning outcomes of STEM lessons taught through the application of robotics as a kinesthetic experiential pedagogical tool.

In the future, the survey will be conducted with a larger number of STEM teachers and learners to enhance the generality of the results. The expected learning outcomes for other grades of students will be investigated. The results will be verified and validated using other robotic platforms for teaching more STEM lessons to K-12 and college students.

Conflicts of Interest: The author declares no conflict of interest.

Ethics Statements: The study was guided by ethics. The study was conducted following local ethical standards and principles for human subjects.

Appendix A

Teacher's code:

Subject(s) teaches: Math/Science (Circle as appropriate)

Years of experience:

<div align="center">Survey Questionnaires</div>

Q 1: What criteria do you perceive to be used to evaluate the learning outcomes of the lesson(s) you have taught?

Answer:

....

....

....

...

Q 2: Do you expect/perceive different learning outcomes for different grades (6-8) of students? (Circle yes or no)

Answer:

Yes/No

Appendix B

Table A1. Proposed metric for assessing each criterion of learning outcomes of students for robotics-enabled math lessons.

Assessment Criteria (Learning Outcomes)	Assessment/Measurement Metrics/Scales and Methods	Proposed Format to Express the Measurement
Test results	Test scores on selected math topics can be used to assess this criterion. Quizzes/tests can be arranged by concerned teachers. In addition, the Dimension of Success (DoS) observation tool can be used to assess math knowledge and practices [31].	percentage (%) of test scores obtained
Computational thinking ability [2]	Computational thinking can be assessed based on custom-developed specific problem-solving scenarios developed and implemented by the teachers. For example, a specific scenario can be developed where students need to solve a particular problem that reflects students' computational thinking abilities. The teachers can observe the students and assess the computational thinking ability of each student separately or of the team as a whole. The teachers can use a 7-point Likert scale to rate the computation thinking ability subjectively based on observations (see note 1). In addition, the computational thinking can be assessed taking inspiration from the methods proposed by Kong [32].	Subjective rating score (see note 2)
Intrinsic and extrinsic motivation	Intrinsic and extrinsic motivation expressed through students' interest in math and their awareness levels for their math-related careers can be assessed directly using a subjective rating scale (e.g., a 7-point Likert scale) based on observations and interviews with the participating students administered by concerned teachers [4]. In addition, the Intrinsic Motivation Instrument (i.e., Self-Determination [33]) may be used to assess the motivation levels of the students for their career path in math. STEM Career Awareness tool may be used to assess their math-related career awareness levels [34]. The PEAR Institute's Common Instrument Suite Student (CIS-S) survey may be used to assess students' math-related attitudes in terms of math engagement, identity, career interest, and career knowledge and activity participation [35]. The DoS can be used to assess math activity engagement, math practices (inquiry and reflection), and youth development in math [31].	Subjective rating score
Trust in robotics	Trust of students in robotics as a pedagogical tool expressed through students' interest to rely on or to believe in the math-related solutions provided by the robotic system can be assessed directly by concerned teachers using a subjective rating scale (e.g., a 7-point Likert scale) based on observations and interviews with the participating students administered by the concerned teachers [5]. See note 3 for more.	Subjective rating score

Table A1. *Cont.*

Assessment Criteria (Learning Outcomes)	Assessment/Measurement Metrics/Scales and Methods	Proposed Format to Express the Measurement
Engagement in class activities	Work sampling method may be used to assess students' engagement in their robotics-enabled lessons [36]. In this method, the teachers may observe each student separately or the team as a whole after a specified time interval (e.g., after every 5 min) during the class, and mark whether they are engaged in their lessons or not. At the end of the observations, the percentage of total class time the students are engaged (or not engaged) can be determined. This is a probabilistic but quantitative assessment method. The following formula may be used to assess student engagement (E) using work sampling, where O_t is the total number of observations in a class and O_e is the total number of observations in that class when student(s) was/were found engaged. $$E = \frac{O_e}{O_t} \times 100\%$$	Percentage (%) of total class time students are engaged in the lesson
Class attendance and punctuality	Attendance record can be used to assess each student's attendance and punctuality (e.g., timely attendance or late attendance) in the class. Percentage (%) of attendance in a specific time period can be calculated. In addition, percentage of timely or late attendance in a specific time period may also be calculated. The objective is to check if student attendance in regular classes increases after participating in the robotics-enabled lessons or being inspired by the robotics-enabled lessons.	Percentage (%) of attendance
Interpersonal relationship	The teachers can observe the students for their robotics-enabled lessons, identify a few cues related to their interpersonal relationships (e.g., how a student addresses his/her team members, reacts at his/her team members' opinions, etc.), and assess each student or the team as a whole using a 7-point Likert scale for their interpersonal relationships. Alternatively, the assessment may be performed as satisfactory or unsatisfactory. In addition, the CIS-S survey can be used to assess the 21st century skills or the socio-emotional learning (SEL) of the students, e.g., relationships with peer students and teachers [35].	Subjective rating score
Engineering and ICT skills	Tests/quizzes administered by the teachers on students' engineering and ICT skills can be used to assess this criterion.	Percentage (%) of test scores obtained
Life-long learning aspiration	The teachers can observe the students for their robotics-enabled lessons, take interviews of each student to know their future plans and goals about their math learning and applications, and assess each student or the team as a whole using a 7-point Likert scale for their life-long learning aspiration.	Subjective rating score
Hands-on and practical ability	Observations administered by the teachers on students' hands-on practical works during a robotics-enabled lesson can be used to assess this criterion. The teachers can observe the class activities performed by the students and rate the hands-on and practical ability of each student or of the team using a 7-point Likert scale.	Subjective rating score
Lab skills and experiment ability	Observations administered by the teachers on students' lab skills and experiment ability during an experiment conducted by the students as a part of a robotics-enabled lesson can be used to assess this criterion. The teachers can observe the class activities and rate the lab skills and experiment ability of each student or of the team using a 7-point Likert scale.	Subjective rating score
Problem solving ability	Observations administered by the teachers on students' problem solving ability as a part of a robotics-enabled lesson can be used to assess this criterion. Assume, there is a problem related to a real-world situation in a robotics-enabled lesson that the students need to solve using math. The students should identify the problem, formulate the problem and determine the strategies to solve the problem using math knowledge and skills. The teachers can observe the ability of each student or of the team in these efforts and rate their abilities using a 7-point Likert scale. The CIS-S survey can also be used to assess the 21st century skills or the socio-emotional learning (SEL) of students, e.g., problem solving/perseverance [35].	Subjective rating score
Formulation of research strategy	Observations administered by the teachers on students' formulation of research strategy during a robotics-enabled lesson can be used to assess this criterion. Assume, there is a problem in a robotics-enabled lesson that the students need to solve using math. The students should identify the problem, formulate the problem, identify the objective, determine hypotheses and research questions, determine the experimental methods and procedures, and analyze the results with future directions. The teachers can observe the ability of each student or of the team in these efforts and rate their abilities using a 7-point Likert scale.	Subjective rating score

Table A1. *Cont.*

Assessment Criteria (Learning Outcomes)	Assessment/Measurement Metrics/Scales and Methods	Proposed Format to Express the Measurement
Teamwork ability	The youth teamwork skills survey can be used to assess the teamwork ability [37]. In addition, the teachers can observe the students for their robotics-enabled lessons, identify a few cues related to their teamwork ability (e.g., how the students split the entire activities of the lesson and assign them to different team members of the team), and assess each student or the team as a whole using a 7-point Likert scale for their teamwork ability.	Subjective rating score
Cognitive workload in learning	NASA TLX can be administered by the teachers on the participating students at the end of each robotics-enabled lesson [28]. Note that the least cognitive workload is the best [29].	Percentage (%) total cognitive workload
Adapting to new situations and changes	The teachers can observe the students for their robotics-enabled lessons, identify a few cues relevant to adapting to new situations and changes (e.g., whether a student can adjust if he/she is transferred to a new team or if a sudden change occurs in the lesson activities), and assess each student or the team as a whole using a 7-point Likert scale for their ability to adapt with new situations and changes.	Subjective rating score
Respect for diversity and multiculturality	The teachers can observe the students for their robotics-enabled lessons, identify a few cues relevant to respect for diversity and multiculturality (e.g., whether a student can adjust with another team member who has different nationality, color, ethnicity, food habits, etc.), and assess each student or the team as a whole using a 7-point Likert scale for their respect for diversity and multiculturality.	Subjective rating score
Professional ethics	The teachers can observe the students for their robotics-enabled lessons, identify a few ethical cues relevant to the class events (e.g., whether a student captures and records true data and does not manipulate the data) and assess each student or the team as a whole using a 7-point Likert scale for their professional ethics.	Subjective rating score
Troubleshooting and contingency	The teachers can observe the students for their robotics-enabled lessons, identify a few cues related to troubleshooting and contingency (e.g., how a student or a team troubleshoots in case the robotics-based experimental system does not work temporarily), and assess each student or the team as a whole using a 7-point Likert scale for their ability for troubleshooting and contingency.	Subjective rating score
Interdisciplinary/ multidisciplinary abilities	The teachers can observe the students for their robotics-enabled lessons and assess each student or the team as a whole using a 7-point Likert scale for their ability to learn and use interdisciplinary and multidisciplinary knowledge and skills (e.g., math content knowledge combined with engineering and computer programming skills to solve a math problem).	Subjective rating score
Reflexive analysis	The teachers can observe the students for their robotics-enabled lessons, take their interviews, and assess each student or the team as a whole using a 7-point Likert scale for their ability to summarize what they learn during the lesson, identify their limitations and develop action plans for improvements in the next lessons.	Subjective rating score
Critical thinking ability	The teachers can observe the students for their robotics-enabled lessons, identify a few cues related to their critical thinking ability (e.g., how the students analyze and compare different alternative possibilities of experimental procedures based on prior findings), and assess each student or the team as a whole using a 7-point Likert scale for their critical thinking ability. In addition, the CIS-S survey can be used to assess the 21st century skills or the socio-emotional learning (SEL) of the students, e.g., critical thinking [35].	Subjective rating score
Decision making ability	The teachers can observe the students for their robotics-enabled lessons, identify a few cues related to their decision-making ability (e.g., how the students make a decision based on the experimental findings, and how they decide the next experiments based on prior findings), and assess each student or the team as a whole using a 7-point Likert scale for their decision-making ability. In addition, the DORA tool can be used to assess reasoning and decision-making abilities of the students [38].	Subjective rating score
Creativity and innovation	The teachers can observe the students for their robotics-enabled lessons, identify a few cues related to their creativity and innovation (e.g., how the students propose a new configuration of the robotic device to solve a particular math problem), and assess each student or the team as a whole using a 7-point Likert scale for their creativity and innovation. In addition, the creativity and innovation can be assessed by the approach proposed by Barbot, Besançon, and Lubart [39].	Subjective rating score

Table A1. *Cont.*

Assessment Criteria (Learning Outcomes)	Assessment/Measurement Metrics/Scales and Methods	Proposed Format to Express the Measurement
Entrepreneurial ability	The students build a robotic device and verify its suitability to learn math and solve real-world problems using math. Such building practices may inculcate entrepreneurial aspiration in the students, which may direct them towards starting a new business initiative to market their ideas and develop new business ventures in the future. The teachers can observe the students for their robotics-enabled lessons, take interviews of each student to know their business plans if any, and assess each student or the team as a whole using a 7-point Likert scale for their entrepreneurial aspiration or ability. In addition, the entrepreneurial ability of the students can be assessed taking inspiration from the methods proposed by Bejinaru [40], and Coduras, Alvarez and Ruiz [41].	Subjective rating score
Communication skills	The teachers can observe the students for their robotics-enabled lessons, identify a few cues related to their communication skills (e.g., how the students communicate the findings of the experiments during their robotics-enabled lessons to their team leader, teachers and each team member), and assess each student or the team as a whole using a 7-point Likert scale for their communication skills. In addition, the CIS-S survey can be used to assess the 21st century skills or the socio-emotional learning (SEL) of the students, e.g., communication skills [35].	Subjective rating score
Leadership ability	Based on specific tasks and scenarios during students' engagement with the robotics-enabled lesson, the surveys proposed by Mazzetto [42] and Chapman and Giri [43] can be used to assess leadership skills of the students. Alternatively, the teachers can observe the students for their robotics-enabled lessons, identify a few cues related to their leadership ability (e.g., how the students decide their leader for a lesson, how the leader directs the team members towards the goal of the lesson, and how the student members follow the directions of the leader), and assess each student or the team as a whole using a 7-point Likert scale for their leadership ability.	Subjective rating score
Organizational and planning ability	The teachers can observe the students for their robotics-enabled lessons, identify a few issues related to organization and planning of the robotics-enabled lesson (e.g., how the students split the responsibility of each team member and determine and ensure the required resources for each member in each step/phase of the entire lesson), and assess each student or the team as a whole using a 7-point Likert scale for their organizational and planning ability.	Subjective rating score
Social responsibility	The teachers can observe the students for their robotics-enabled lessons, identify a few social cues relevant to the class events (e.g., whether a student wishes another student in his/her birthday that falls on the day of a robotics-enabled lesson, or how a student feels if another student of the team is known to be sick), and assess each student or the team as a whole using a 7-point Likert scale for their social responsibility.	Subjective rating score

Note 1: In the 7-point Likert scale, −3 is the least or worst, 0 is the neutral, and +3 is the highest or the best response. Note 2: The subjective rating score is expressed as a score value between −3 and +3 with a possible difference of |1| between two adjacent scores. Note 3: For some learning outcomes, in addition to the proposed assessment metrics, the assessment may be qualitatively performed as satisfactory or unsatisfactory. Furthermore, teachers can qualitatively assess each outcome and prepare a short qualitative report on each outcome criterion. These can be cross-checked/triangulated with the proposed quantitative metrics under mixed method analyses [2].

Table A2. The mean learning outcomes for different learning outcome criteria for robotics-enabled math lessons for different schools.

Learning Outcome Criteria	Schools																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Test results (see note 4)	93	94	98	92	97	90	91	91	94	89	99	95	96	92	90	97	88	94	99	93
Computational thinking ability	2.29	2.46	2.87	2.33	2.49	2.65	2.27	2.78	2.61	2.82	2.66	2.57	2.49	2.77	2.84	2.19	2.38	2.73	2.92	2.68
Intrinsic and extrinsic motivation	2.67	2.63	2.68	2.45	2.69	2.56	2.46	2.59	2.42	2.73	2.57	2.46	2.74	2.81	2.38	2.54	2.78	2.29	2.85	2.96
Trust in robotics	2.82	2.39	2.68	2.56	2.71	2.36	2.80	2.53	2.44	2.72	2.81	2.75	2.70	2.88	2.52	2.26	2.66	2.42	2.61	2.63
Engagement	84	91	97	89	93	96	95	99	88	89	92	90	93	99	99	86	95	92	88	89
Class attendance and punctuality	96	96	96	95	99	92	94	100	98	98	100	100	100	99	99	96	100	90	96	98
Interpersonal relationship	2.77	2.49	2.56	2.42	2.54	2.76	2.33	2.52	2.48	2.71	2.58	2.39	2.44	2.81	2.67	2.28	2.59	2.64	2.88	2.51
Engineering and ICT skills	96	99	96	98	93	95	98	93	96	100	93	100	99	97	89	100	94	97	100	98
Life-long learning aspiration	2.33	2.40	2.73	2.74	2.38	2.49	2.56	2.62	2.58	2.77	2.54	2.36	2.67	2.69	2.55	2.82	2.64	2.51	2.74	2.45
Hands-on and practical ability	2.56	2.73	2.71	2.62	2.33	2.49	2.89	2.76	2.55	2.64	2.69	2.52	2.60	2.72	2.67	2.66	2.72	2.59	2.48	2.65
Lab skills and experiment ability	2.22	2.46	2.82	2.39	2.68	2.29	2.48	2.63	2.54	2.78	2.61	2.53	2.44	2.46	2.67	2.35	2.48	2.72	2.62	2.66
Problem solving ability	2.54	2.56	2.75	2.66	2.54	2.78	2.46	2.80	2.48	2.43	2.62	2.50	2.83	2.61	2.45	2.37	2.39	2.78	2.29	2.59
Formulation of research strategy	2.36	2.54	2.74	2.39	2.44	2.36	2.47	2.67	2.73	2.72	2.65	2.50	2.43	2.79	2.62	2.37	2.56	2.52	2.76	2.28
Teamwork ability	2.53	2.75	2.73	2.69	2.45	2.77	2.56	2.87	2.54	2.67	2.73	2.54	2.63	2.75	2.69	2.69	2.74	2.67	2.47	2.26
Cognitive workload	23	26	19	31	25	23	28	17	33	12	22	25	28	18	19	27	16	30	19	20
Adapting to new situations and changes	2.46	2.77	2.29	2.38	2.67	2.64	2.60	2.74	2.53	2.71	2.33	2.52	2.79	2.72	2.67	2.35	2.40	2.51	2.68	2.49
Respect for diversity and multiculturality	2.28	2.41	2.76	2.38	2.46	2.66	2.25	2.46	2.66	2.81	2.59	2.54	2.50	2.79	2.81	2.27	2.37	2.65	2.63	2.87
Professional ethics	3.00	2.95	2.87	2.68	2.50	2.66	2.55	2.80	2.66	2.56	3.00	2.90	2.95	3.00	3.00	2.86	2.56	2.75	2.88	2.43
Troubleshooting and contingency	2.56	2.44	2.65	2.32	2.57	2.63	2.71	2.62	2.74	2.48	2.69	2.54	2.44	2.62	2.26	2.75	2.42	2.61	2.78	2.81
Interdisciplinary/ multidisciplinary abilities	2.91	2.52	2.28	2.79	2.42	2.33	2.48	2.30	2.59	2.77	2.63	2.51	2.82	2.38	2.61	2.75	2.39	2.54	2.83	2.47

Table A2. *Cont.*

Learning Outcome Criteria	Schools																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Reflexive analysis	2.45	2.49	2.73	2.24	2.35	2.54	2.84	2.63	2.36	2.29	2.48	2.83	2.27	2.54	2.56	2.47	2.66	2.77	2.91	2.89
Critical thinking ability	2.31	2.49	2.69	2.60	2.48	2.63	2.45	2.39	2.78	2.57	2.64	2.65	2.91	2.43	2.63	2.67	2.71	2.29	2.70	2.22
Decision making ability	2.75	2.78	2.75	2.50	2.68	2.60	2.81	2.44	2.29	2.34	2.63	2.33	2.46	2.72	2.80	2.46	2.50	2.48	2.88	2.45
Creativity and innovation	2.82	2.67	2.91	2.30	2.43	2.66	2.78	2.73	2.41	2.85	2.63	2.47	2.69	2.37	2.64	2.62	2.43	2.79	2.90	2.86
Entrepreneurial ability	2.54	2.45	2.84	2.39	2.47	2.61	2.29	2.73	2.60	2.80	2.56	2.52	2.50	2.38	2.83	2.71	2.78	2.69	2.82	2.58
Communication skills	2.12	2.87	2.28	2.39	2.90	2.85	2.22	2.62	2.60	2.67	2.60	2.49	2.58	2.29	2.81	2.33	2.45	2.74	2.35	2.42
Leadership ability	2.31	2.49	2.72	2.36	2.62	2.69	2.28	2.66	2.63	2.56	2.37	2.58	2.57	2.72	2.80	2.59	2.33	2.44	2.08	2.66
Organizational and planning ability	2.28	2.56	2.67	2.42	2.42	2.68	2.33	2.58	2.89	2.83	2.63	2.55	2.46	2.70	2.23	2.17	2.57	2.42	2.65	2.75
Social responsibility	2.28	2.72	2.29	2.34	2.48	2.60	2.22	2.19	2.56	2.65	2.62	2.53	2.62	2.75	2.81	2.56	2.39	2.72	2.91	2.62

Note 4: The mean score of test results for School#1 was 93. Its meaning is as follows: assume 10 students participated in the robotics-enabled math lesson. The teacher determined the mean of the test scores obtained by all of the 10 students in the math test after the math lesson, and found 93 (rounded) as the mean score. Other scores were calculated in the similar way.

Table A3. The mean learning outcomes for different learning outcome criteria for robotics-enabled science lessons for different schools

Learning Outcome Criteria	Schools																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Test results	94	96	99	94	99	94	97	96	95	93	99	98	97	95	94	98	93	96	98	99
Imagination ability	2.33	2.47	2.88	2.45	2.54	2.69	2.43	2.93	2.69	2.88	2.75	2.78	2.64	2.79	2.86	2.65	2.61	2.78	2.83	2.71
Intrinsic and extrinsic motivation	2.69	2.67	2.72	2.48	2.73	2.59	2.48	2.64	2.47	2.76	2.59	2.48	2.76	2.86	2.39	2.56	2.82	2.43	2.76	2.97
Trust in robotics	2.85	2.67	2.76	2.59	2.78	2.38	2.85	2.57	2.48	2.79	2.85	2.77	2.71	2.89	2.57	2.28	2.69	2.41	2.66	2.65
Engagement	86	93	98	90	94	98	97	99	90	91	93	92	94	99	99	88	98	94	89	90
Class attendance and punctuality	99	97	98	98	100	93	95	100	99	99	100	100	98	98	99	97	100	96	97	99
Interpersonal relationship	2.79	2.53	2.59	2.45	2.58	2.79	2.36	2.56	2.49	2.75	2.63	2.43	2.46	2.85	2.69	2.33	2.63	2.67	2.89	2.54
Engineering and ICT skills	100	99	98	99	96	97	99	96	100	99	96	99	100	98	92	99	98	100	99	100
Life-long learning aspiration	2.36	2.46	2.74	2.77	2.39	2.54	2.58	2.65	2.59	2.79	2.58	2.39	2.68	2.72	2.58	2.83	2.67	2.55	2.76	2.52
Hands-on and practical ability	2.58	2.77	2.75	2.66	2.36	2.50	2.93	2.78	2.59	2.66	2.72	2.54	2.68	2.79	2.68	2.67	2.83	2.65	2.49	2.68

Table A3. *Cont.*

Learning Outcome Criteria	Schools																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Lab skills and experiment ability	2.26	2.49	2.87	2.43	2.69	2.35	2.49	2.66	2.59	2.79	2.65	2.57	2.47	2.49	2.68	2.39	2.65	2.78	2.65	2.69
Problem solving ability	2.63	2.58	2.78	2.67	2.57	2.79	2.47	2.84	2.49	2.47	2.69	2.62	2.74	2.65	2.54	2.42	2.53	2.79	2.35	2.64
Formulation of research strategy	2.39	2.55	2.78	2.43	2.46	2.39	2.49	2.68	2.68	2.83	2.74	2.68	2.47	2.82	2.68	2.39	2.59	2.56	2.78	2.34
Teamwork ability	2.56	2.81	2.79	2.76	2.47	2.79	2.57	2.89	2.58	2.69	2.76	2.57	2.68	2.77	2.83	2.73	2.76	2.68	2.49	2.54
Cognitive workload	21	23	18	22	21	20	24	18	25	18	20	13	22	14	16	23	15	19	16	18
Adapting to new situations and changes	2.48	2.78	2.34	2.39	2.72	2.66	2.66	2.77	2.59	2.82	2.42	2.61	2.82	2.78	2.68	2.46	2.48	2.64	2.69	2.53
Respect for diversity and multiculturality	2.35	2.49	2.77	2.39	2.47	2.68	2.34	2.49	2.73	2.78	2.63	2.55	2.54	2.78	2.88	2.56	2.67	2.68	2.67	2.85
Professional ethics	2.95	2.99	2.88	2.75	2.58	2.69	2.60	2.84	2.68	2.58	2.90	2.92	2.98	2.98	2.87	2.88	2.59	2.79	2.65	2.89
Troubleshooting and contingency	2.62	2.46	2.68	2.62	2.78	2.64	2.74	2.66	2.79	2.53	2.76	2.68	2.46	2.67	2.73	2.76	2.56	2.66	2.79	2.84
Interdisciplinary/ multidisciplinary abilities	2.95	2.62	2.56	2.83	2.65	2.78	2.56	2.34	2.65	2.78	2.69	2.57	2.88	2.46	2.68	2.78	2.48	2.73	2.67	2.58
Reflexive analysis	2.49	2.50	2.77	2.29	2.38	2.58	2.85	2.64	2.39	2.46	2.49	2.85	2.44	2.66	2.60	2.56	2.68	2.78	2.95	2.92
Critical thinking ability	2.34	2.54	2.72	2.65	2.49	2.66	2.48	2.46	2.79	2.63	2.68	2.81	2.87	2.49	2.68	2.68	2.73	2.45	2.76	2.49
Decision making ability	2.76	2.79	2.78	2.57	2.69	2.66	2.88	2.49	2.56	2.39	2.67	2.39	2.49	2.78	2.89	2.61	2.58	2.60	2.91	2.48
Creativity and innovation	2.89	2.78	2.92	2.45	2.48	2.69	2.79	2.77	2.48	2.89	2.65	2.49	2.74	2.56	2.66	2.67	2.49	2.83	2.93	2.89
Entrepreneurial ability	2.56	2.49	2.85	2.43	2.48	2.65	2.33	2.75	2.65	2.85	2.58	2.55	2.56	2.39	2.86	2.76	2.79	2.78	2.85	2.62
Communication skills	2.34	2.89	2.54	2.45	2.91	2.86	2.57	2.68	2.67	2.71	2.65	2.50	2.59	2.33	2.82	2.53	2.48	2.78	2.39	2.45
Leadership ability	2.39	2.54	2.75	2.43	2.63	2.70	2.47	2.68	2.64	2.58	2.64	2.62	2.73	2.74	2.82	2.67	2.38	2.46	2.23	2.75
Organizational and planning ability	2.51	2.66	2.68	2.59	2.54	2.69	2.56	2.64	2.92	2.86	2.66	2.56	2.49	2.72	2.46	2.36	2.68	2.54	2.78	2.79
Social responsibility	2.63	2.78	2.56	2.59	2.52	2.66	2.37	2.45	2.58	2.66	2.67	2.58	2.67	2.78	2.86	2.68	2.42	2.77	2.94	2.72

References

1. Rahman, S.M.M.; Krishnan, V.J.; Kapila, V. Optimizing a teacher professional development program for teaching STEM with robotics through design-based research. In Proceedings of the 2018 ASEE Annual Conference & Exposition, Salt Lake City, UT, USA, 24–27 June 2018; pp. 1–20.
2. Rahman, S.M.M.; Chacko, S.M.; Rajguru, S.B.; Kapila, V. Determining prerequisites for middle school students to participate in robotics-based STEM lessons: A computational thinking approach. In Proceedings of the 2018 ASEE Annual Conference & Exposition, Salt Lake City, UT, USA, 24–27 June 2018; pp. 1–27.
3. Mallik, A.; Rahman, S.M.M.; Rajguru, S.B.; Kapila, V. Examining the variations in the TPACK framework for teaching robotics-aided STEM lessons of varying difficulty. In Proceedings of the 2018 ASEE Annual Conference & Exposition, Salt Lake City, UT, USA, 24–27 June 2018; pp. 1–23.
4. Rahman, S.M.M.; Krishnan, V.J.; Kapila, V. Exploring the dynamic nature of TPACK framework in teaching STEM using robotics in middle school classrooms. In Proceedings of the 2017 ASEE Annual Conference & Exposition, Columbus, OH, USA, 25–28 June 2017; pp. 1–29.
5. Rahman, S.M.M.; Chacko, S.M.; Kapila, V. Building trust in robots in robotics-focused STEM education under TPACK framework in middle schools. In Proceedings of the 2017 ASEE Annual Conference & Exposition, Columbus, OH, USA, 25–28 June 2017; pp. 1–25.
6. Rahman, S.M.M.; Kapila, V. A systems approach to analyzing design-based research in robotics-focused middle school STEM lessons through cognitive apprenticeship. In Proceedings of the 2017 ASEE Annual Conference & Exposition, Columbus, OH, USA, 25–28 June 2017; pp. 1–25.
7. Chen, N.S.; Quadir, B.; Teng, D.C. Integrating book, digital content and robot for enhancing elementary school students' learning of English. *Australas. J. Educ. Technol.* **2011**, *27*, 546–561. [CrossRef]
8. Mosley, P.; Kline, R. Engaging students: A framework using LEGO robotics to teach problem solving. *Inf. Technol. Learn. Perform. J.* **2006**, *24*, 39–45.
9. Whitman, L.; Witherspoon, T. Using LEGOs to interest high school students and improve K12 STEM education. In Proceedings of the ASEE/IEEE Frontiers in Education Conference, Westminster, CO, USA, 5–8 November 2003; p. F3A6-10.
10. Belpaeme, T.; Kennedy, J.; Ramachandran, A.; Scassellati, B.; Tanaka, F. Social robots for education: A review. *Sci. Robot.* **2018**, *3*, eaat5954. [CrossRef] [PubMed]
11. Toh, L.P.E.; Causo, A.; Tzuo, P.; Chen, I.; Yeo, S.H. A review on the use of robots in education and young children. *J. Educ. Technol. Soc.* **2016**, *19*, 148–163.
12. Barreto, F.; Benitti, V. Exploring the educational potential of robotics in schools: A systematic review. *Comput. Educ.* **2012**, *58*, 978–988.
13. Danahy, E.; Wang, E.; Brockman, J.; Carberry, A.; Shapiro, B.; Rogers, C.B. LEGO-based robotics in higher education: 15 years of student creativity. *Int. J. Adv. Robot. Syst.* **2014**, *11*, 27. [CrossRef]
14. Rahman, S.M.M. Instructing a mechatronics course aligning with TPACK framework. In Proceedings of the 2019 ASEE Annual Conference & Exposition, Tampa, FL, USA, 15–19 June 2019.
15. Rahman, S.M.M. Instruction design of a mechatronics course based on closed-loop 7E model refined with DBR method. In Proceedings of the 2019 ASEE Annual Conference & Exposition, Tampa, FL, USA, 15–19 June 2019.
16. Rahman, S.M.M. Comparative experiential learning of mechanical engineering concepts through the usage of robot as a kinesthetic learning tool. In Proceedings of the 2019 ASEE Annual Conference & Exposition, Tampa, FL, USA, 15–19 June 2019.
17. Erikson, M.G.; Erikson, M. Learning outcomes and critical thinking–good intentions in conflict. *Stud. High. Educ.* **2018**, *44*, 2293–2303. [CrossRef]
18. Brooks, S.; Dobbins, K.; Scott, J.J.A.; Rawlinson, M.; Norman, R.I. Learning about learning outcomes: The student perspective. *Teach. High. Educ.* **2014**, *19*, 721–733. [CrossRef]
19. Melguizo, T.; Coates, H. The value of assessing higher education student learning outcomes. *AERA Open* **2017**, *3*. [CrossRef]
20. Prøitz, T.S. Learning outcomes: "What are they? *Who defines them? When and where are they defined?"* *Educ. Assess. Eval. Account.* **2010**, *22*, 119–137.
21. Farquharson, K. Regulating sociology: Threshold learning outcomes and institutional isomorphism. *J. Sociol.* **2013**, *49*, 486–500. [CrossRef]
22. Watson, P. The role and integration of learning outcomes into the educational process. *Act. Learn. High. Educ.* **2002**, *3*, 205–219. [CrossRef]
23. Smith, B.W.; Zhou, Y. Assessment of learning outcomes: The example of spatial analysis at Bowling Green State University. *Int. Res. Geogr. Environ. Educ.* **2005**, *14*, 211–216. [CrossRef]
24. Oliver, B.; Tucker, B.; Gupta, R.; Yeo, S. eVALUate: An evaluation instrument for measuring students' perceptions of their engagement and learning outcomes. *Assess. Eval. High. Educ.* **2008**, *33*, 619–630. [CrossRef]
25. Svanström, M.; Lozano-García, F.J.; Rowe, D. Learning outcomes for sustainable development in higher education. *Int. J. Sustain. High. Educ.* **2008**, *9*, 339–351. [CrossRef]
26. Shephard, K. Higher education for sustainability: Seeking affective learning outcomes. *Int. J. Sustain. High. Educ.* **2008**, *9*, 87–98. [CrossRef]

27. Rahman, S.M.M.; Ikeura, R. Calibrating intuitive and natural human-robot interaction and performance for power-assisted heavy object manipulation using cognition-based intelligent admittance control schemes. *Int. J. Adv. Robot. Syst.* **2018**, *15*. [CrossRef]

28. Rahman, S.M.M.; Ikeura, R. Cognition-based variable admittance control for active compliance in flexible manipulation of heavy objects with a power assist robotic system. *Robot. Biomim.* **2018**, *5*, 1–25.

29. de Jong, T. Cognitive load theory, educational research, and instructional design: Some food for thought. *Instruct. Sci.* **2010**, *38*, 105–134. [CrossRef]

30. Leite, A.; Soares, D.; Sousa, H.; Vidal, D.; Dinis, M.; Dias, D. For a healthy (and) higher education: Evidences from learning outcomes in health sciences. *Educ. Sci.* **2020**, *10*, 168. [CrossRef]

31. DoS. Available online: https://www.thepearinstitute.org/dimensions-of-success (accessed on 18 December 2020).

32. Kong, S. Components and methods of evaluating computational thinking for fostering creative problem-solvers in senior primary school education. In *Computational Thinking Education*; Kong, S., Abelson, H., Eds.; Springer: Singapore, 2019. [CrossRef]

33. Deci, E.L.; Eghrari, H.; Patrick, B.C.; Leone, D. Facilitating internalization: The self-determination theory perspective. *J. Personal.* **1994**, *62*, 119–142. [CrossRef]

34. Kier, M.W.; Blanchard, M.R.; Osborne, J.W.; Albert, J.L. The development of the STEM career interest survey (STEM-CIS). *Res. Sci. Educ.* **2014**, *44*, 461–481. [CrossRef]

35. CIS-S. Available online: http://www.pearweb.org/atis/tools/common-instrument-suite-student-cis-s-survey (accessed on 18 December 2020).

36. Robinson, M. Work sampling: Methodological advances and new applications. *Hum. Factors Ergon. Manuf. Serv. Ind.* **2009**, *20*, 42–60. [CrossRef]

37. Available online: https://www.informalscience.org/youth-teamwork-skills-survey-manual-and-survey (accessed on 18 December 2020).

38. Algozzine, A.; Newton, J.; Horner, R.; Todd, A.; Algozzine, K. Development and technical characteristics of a team decision-making assessment tool: Decision observation, recording, and analysis (DORA). *J. Psychoeduc. Assess.* **2012**, *30*, 237–249. [CrossRef]

39. Barbot, B.; Besançon, M.; Lubart, T. Assessing creativity in the classroom. *Open Educ. J.* **2011**, *4*, 58–66. [CrossRef]

40. Bejinaru, R. Assessing students' entrepreneurial skills needed in the knowledge economy. *Manag. Mark. Chall. Knowl. Soc.* **2018**, *13*, 1119–1132. [CrossRef]

41. Coduras, A.; Alvarez, J.; Ruiz, J. Measuring readiness for entrepreneurship: An information tool proposal. *J. Innov. Knowl.* **2016**, *1*, 99–108. [CrossRef]

42. Mazzetto, S. A practical, multidisciplinary approach for assessing leadership in project management education. *J. Appl. Res. High. Educ.* **2019**, *11*, 50–65. [CrossRef]

43. Chapman, A.; Giri, P. Learning to lead: Tools for self-assessment of leadership skills and styles. In *Why Hospitals Fail*; Godbole, P., Burke, D., Aylott, J., Eds.; Springer: Cham, Switzerland, 2017. [CrossRef]

44. NGSS. *Next Generation Science Standards (NGSS): For States, by States*; The National Academies Press: Washington, DC, USA, 2013; Available online: http://www.nextgenscience.org/ (accessed on 20 December 2020).

45. Elen, J.; Clarebout, G.; Leonard, R.; Lowyck, J. Student-centred and teacher-centred learning environments: What students think. *Teach. High. Educ.* **2007**, *12*, 105–117. [CrossRef]

46. Chan, C.C.; Tsui, M.S.; Chan, M.Y.C.; Hong, J.H. Applying the structure of the observed learning outcomes (SOLO) taxonomy on student's learning outcomes: An empirical study. *Assess. Eval. High. Educ.* **2002**, *27*, 511–527. [CrossRef]

47. Koretsky, M.; Keeler, J.; Ivanovitch, J.; Cao, Y. The role of pedagogical tools in active learning: A case for sense-making. *Int. J. STEM Educ.* **2018**, *5*, 18. [CrossRef] [PubMed]

48. Garbett, D.L. Assignments as a pedagogical tool in learning to teach science: A case study. *J. Early Child. Teach. Educ.* **2007**, *28*, 381–392. [CrossRef]

49. Ahmed, A.; Clark-Jeavons, A.; Oldknow, A. How can teaching aids improve the quality of mathematics education. *Educ. Stud. Math.* **2004**, *56*, 313–328. [CrossRef]

50. Geiger, T.; Amrein-Beardsley, A. Student perception surveys for K-12 teacher evaluation in the United States: A survey of surveys. *Cogent Educ.* **2019**, *6*, 1602943. [CrossRef]

51. Sturtevant, H.; Wheeler, L. The STEM Faculty Instructional Barriers and Identity Survey (FIBIS): Development and exploratory results. *Int. J. STEM Educ.* **2019**, *6*, 1–22. [CrossRef]

52. Jones, N.D.; Brownell, M.T. Examining the use of classroom observations in the evaluation of special education teachers. *Assess. Eff. Interv.* **2014**, *39*, 112–124. [CrossRef]

53. Rizi, C.; Najafipour, M.; Haghani, F.; Dehghan, S. The effect of the using the brainstorming method on the academic achievement of students in grade five in Tehran elementary schools. *Procedia Soc. Behav. Sci.* **2013**, *83*, 230–233. [CrossRef]

54. Ritzhaupt, A.D.; Dawson, K.; Cavanaugh, C. An investigation of factors influencing student use of technology in K-12 classrooms using path analysis. *J. Educ. Comput. Res.* **2012**, *46*, 229–254. [CrossRef]

55. Batdi, V.; Talan, T.; Semerci, C. Meta-analytic and meta-thematic analysis of STEM education. *Int. J. Educ. Math. Sci. Technol.* **2019**, *7*, 382–399.

56. Gao, X.; Li, P.; Shen, J.; Sun, H. Reviewing assessment of student learning in interdisciplinary STEM education. *Int. J. STEM Educ.* **2020**, *7*, 1–14. [CrossRef]

57. Hartikainen, S.; Rintala, H.; Pylväs, L.; Nokelainen, P. The concept of active learning and the measurement of learning outcomes: A review of research in engineering higher education. *Educ. Sci.* **2019**, *9*, 276. [CrossRef]

58. Miskioğlu, E.E.; Asare, P. Critically thinking about engineering through kinesthetic experiential learning. In Proceedings of the 2017 IEEE Frontiers in Education Conference (FIE), Indianapolis, IN, USA, 18–21 October 2017; pp. 1–3.

59. Emerson, L.; MacKay, B. A comparison between paper-based and online learning in higher education. *Br. J. Educ. Technol.* **2011**, *42*, 727–735. [CrossRef]

60. Collins, A. Cognitive apprenticeship and instructional technology. In *Educational Values and Cognitive Instruction: Implications for Reform*; Routledge: New York, NY, USA, 1991; pp. 121–138.

61. Capraro, R.; Slough, S. Why PBL? Why STEM? Why now? An Introduction to STEM Project-Based Learning. In *STEM Project-Based Learning*; Capraro, R., Capraro, M., Morgan, J., Eds.; Sense Publishers: Rotterdam, The Netherlands, 2013. [CrossRef]

62. Dringenberg, E.; Wertz, R.; Purzer, S.; Strobel, J. Development of the science and engineering classroom learning observation protocol. In Proceedings of the 2012 ASEE Annual Conference & Exposition, San Antonio, TX, USA, 10–13 June 2012.

63. Michelsen, C. IBSME—Inquiry-based science and mathematics education. *MONA-Matematik-Og Naturfagsdidaktik* **2011**, *6*, 72–77.

64. Anderson, J.R.; Reder, L.M.; Simon, H.A. Situated learning and education. *Educ. Res.* **1996**, *25*, 5–11. [CrossRef]