Development of Performance Assessment on Science Practicum Integrated with Automated Feedback to Measure Scientific Attitude in University: a Case Study in Indonesia

Sudirman [1,2], Agus Ramdani [1], Aris Doyan [1], Yunita Arian Sani Anwar [1], Joni Rokhmat [1], AA. Sukarso [1]

[1] *University of Mataram*
Jl. Majapahit No 62 Mataram, Nusa Tenggara Barat, Indonesia

[2] *University of Qamarul Huda Badaruddin Bagu*
Jl. Badaruddin Bagu Pringgarata, Central Lombok, West Nusa Tenggara, Indonesia

DOI: 10.22178/pos.99-7

LCC Subject Category: L7-991

Received 04.11.2023
Accepted 29.12.2023
Published online 31.12.2023

Corresponding Author:
Sudirman
sudirmansiddiq@gmail.com

Abstract. The purpose of this research is to know the responses of respondents and university types on the development of scientific attitude instruments that are integrated with automatic feedback on science practicum performance assessments at four universities in Indonesia. Using a case study approach and cluster random sampling procedures, the data were examined with a Kruskal-Wallis test. Questionnaires, interviews at the management level, and observations directly were used to collect data. The research sample included three hundred seventy-six undergraduates and teachers in college at four different universities. The findings reveal no performance assessment of scientific attitude since there was no standard instrument. The assessment results of private colleges were similar to those of public colleges ($p > 0.05$) in measuring scientific attitude integrated with real-time feedback, implying that neither public nor private universities have conducted scientific attitude of performance assessments with real-time feedback on science practicum. The analysis statistical results were not significant between undergraduates and teachers in college regarding the measurement of scientific attitude on performance assessments integrated with real-time feedback in science practicum ($p > 0.05$), implying that undergraduates have the same experience as lecturers in assessing scientific attitude on performance assessment with feedback in science practicum. This study suggests developing scientific attitude measurement on performance assessment that provides real-time science practicum feedback.

Keywords: Performance Assessment; Science Education; Scientific Attitude.

INTRODUCTION

The term "assessment" refers to systematically collecting qualitative and quantitative data concerning the achievements and characteristics of a science program [1]. According to the information provided in reference [4], assessment is defined as the methodical gathering of information about students, particularly concerning their activities and knowledge. The professors can monitor the students, evaluate their skills and knowledge, and examine the products that the students generate while they are collecting information. Assessment activities consider not only the characteristics of the students being evaluated but also the characteristics of the instructional methods, the curriculum, the facilities, and the administration of the teaching institution. "Assessment instruments" refers to various official and informal measures or processes utilised to gather information about students. Written examinations, oral examinations, observation sheets, interview guides, and homework assignments are examples of tests that may be included [16].

The implementation of student learning activities that integrate tasks that involve advanced cognitive capacities and the integration of broad knowledge by contextual learning objectives is an example of authentic assessment. When the evaluation of learning goes beyond the classroom bounds and represents something significant for the students, then the assessment is considered to have achieved authenticity. To conduct an authentic review, you need to possess the exact skills and capabilities required to execute the work properly. An accurate assessment is advantageous in several different ways. First, it requires students to come up with their responses instead of choosing from a list of already established possibilities. The second benefit is that it encourages higher-level thinking and fundamental skills. Thirdly, it provides an in-depth analysis of a thorough undertaking or assignment.

Last but not least, it is included in the instructional process. Lastly, it uses examples of students' work, such as a portfolio. Sixthly, it is founded on well-defined standards for students to follow. In the seventh place, it makes it possible for various perspectives to arise. Eighthly, it shares a lot of similarities with learning in the classroom. The final benefit is that it teaches students how to evaluate their work [16].

The evaluation method utilised to ascertain the outcome of the learning process is known as authentic assessment. Authentic assessment is relevant to the scientific method in education, as described by the standards of the Curriculum 2013, which may be found here. It does an excellent job of measuring learners' development in their educational achievements. Authentic evaluation aims to assess a person's capabilities in a variety of situations that are representative of real-world situations [2]. Accurate assessment is a sort of assessment that can be used to measure skill ability, and performance assessment is included in this category. Performance evaluation is a component of the educational process in the field of science, which is one of the subjects taught in an academic setting.

The systematic gathering of information about the natural world is what we mean when discussing science. Naturally, it is frequently referred to as Natural Science when it is addressed in the context of education. Not only does the process of acquiring knowledge in the form of facts, concepts, or principles fall under the purview of science, but so does the act of exploring and discovering new things. Consequently, the teaching of science in schools ought to be carried out in a manner that is both systematic and founded on evidence. Students ought to be equipped with skills that are specifically tailored for the process of learning science. Performance assessment is the appraisal of learners' performance and achievements in their tasks, as shown by their actions or presentations as assigned by teachers [10]. Performance assessment is also known as performance evaluation. Students are required to demonstrate their knowledge and abilities through various concrete activities or performances to be evaluated on their performance. As part of the Performance Test, students must demonstrate their ability to accomplish tasks based on the actual world and appropriately reflect their overall performance. Some examples of these jobs are the preparation of tools, the utilisation of tools, the assembling of tools, the writing of data, the analysis of data, the summarising of data, and the reporting of data. Students' aptitude, comprehension, ability to apply information and skills, execution, and proficiency in carrying out tasks are all mainly evaluated during performance evaluations. A collection of knowledge that includes factual facts, conceptual frameworks, and fundamental principles of the scientific area is what we refer to as science. Science is a commodity. Science is a comprehensive concept encompassing scientists' skills and perspectives [21].

The results of several studies that were carried out on educators [5] highlight that there is a deficiency in the utilisation of efficient performance evaluations in the educational process. The findings of this study reveal that during the preparation stage for practicum, there is a noticeable shortfall in the criteria for developing clear and thorough explanations, with a score of 59%. This research investigates the execution of scientific practicum, and the findings of this study imply that this weakness is apparent. In addition, a survey carried out by [17] investigated the difficulties that Yogyakarta City teachers encountered when attempting to execute the assessment component of the curriculum implemented in 2013. The assessment of the abilities possessed by the students was one of the specific issues that was recognised. According to the study's findings, educators generally lack understanding regarding the concept of skills evaluation. Developing a skill evaluation rubric and determining which indicators are most appropriate for evaluating student skills are complicated. One possible explanation for this is that teachers' perceptions and preparedness to consider the practicum method are not as satisfactory as they may be. Schools and other educational institutions should offer teachers extensive training on developing and implementing authentic assessment [3]. This training should be provided to teachers through schools.

The notion of authentic assessment is expansive and open to interpretation, much like the definition of performance evaluation. A description of the various perspectives and focuses adopted when considering authenticity in assessment is included in the first subsection of this introductory section. In addition to providing an overview of the basic meanings of authentic assessment, the description can also be utilised to categorise the various ways the term is understood. The following two subsections provide numerous definitions of accurate assessment, highlighting the multiple perspectives and areas of attention discussed in the previous section. The first of these two subsections concerns generic notions, while the second concerns definitions especially relevant to mathematics in the classroom. The objective has been to select instances illustrative of the many meanings related to the viewpoints and focuses that have been stated. The purpose of the presented definitions is to show and explain the many perspectives and emphases and outline the significance of these perspectives in terms of the differences and similarities that exist in the interpretations of authentic evaluation. This endeavour aims not to identify the fundamental qualities that have been acknowledged but rather to thoroughly embrace all aspects of the several interpretations in exquisite detail. Authentic evaluation is the primary topic of discussion throughout this section. Because activities are the fundamental components and occupy a crucial position in various assessment formats, and they must be judged genuinely for the assessments to be authentic, concepts about assessment tasks are also considered.

The feedback may consist of two components: 1) providing straightforward information concerning the correctness or incorrectness of the response and 2) additional details concerning the student's performance in the test as a whole or concerning a particular answer. Enlarged feedback is more effective in educational contexts and can significantly boost learning [5]. This is according to research that has been conducted. Furthermore, the research by [8] reveals that the effectiveness of elaborated feedback in aiding learning is boosted when adapted to the specific context of each activity. This type of feedback is generally referred to as contextualised feedback. According to the findings of several research studies, three essential components need to be included in input for it to be successful and helpful. First, it ought to provide an accurate evaluation of the student's existing knowledge or capabilities, also known as their actual performance [1]. Second, it should determine the areas in which the student has to improve, also referred to as their desirable performance [4]. In conclusion, it must offer suggestions or direction regarding how the student might achieve the targeted level of performance [18, 2].

Technology-based assessment, namely performance evaluation and diagnostic assessment, has been recognised as a vital instrument in supporting the day-to-day educational process due to the technological revolution that has taken place in education. Students can successfully track their development through such assessments, which are essential sources of information and feedback. Additionally, scientific attitude evaluation can be merged with automatically provided feedback between instructors and university students. Consequently, it is necessary to carry out a study using the distribution of online questionnaires, the conduct of face to face interviews, and the observation of laboratory operations at 4 different colleges, with the following research questions:

1. What is the present status of the four universities on science practicum assessment to scientific attitude?

2. Whether the types of universities and the respondents' responses have a statistically significant effect on the assessment of scientific attitude?

METHOD

Research Design. The research methodology employed in this work adopts a case study approach, incorporating both quantitative and qualitative methods, and utilises the cluster random sampling technique.

Population and Sample. Every state and private university in the West Nusa Tenggara Province was included in this research project's sample. Participants were chosen from four different colleges to obtain the sample, which was collected through cluster or random sampling. Mataram University and Mataram State Islamic University were two public universities included in this group of educational establishments. Additionally, two private universities were included in this group: Universitas Pendidikan Mandalika and Universitas Qamarul Huda Badaruddin Bagu. A total of 376 people are included in the study population, with 261 students from universities and 115 professors. A total of 101 of them are male, while 275 of them are female. The data presented in Table 1 shows that most of the participants are 21 years old. Interviews were conducted with 4 academic deans, 4 study program directors, 4 laboratory heads, and 4 lecturers assistant. The purpose of these interviews was to validate the data findings and enrich the dataset collected from the online survey.

Research Instruments. The instrument is composed of three parts, the first of which is concerned with the characteristics of the participants. These participants include university students, lecturers, gender, age, and the type of university cluster they belong to (public or private). The following section is concerned with providing evidence for evaluating scientific attitudes in science practicum, and it contains a total of eighteen different question items. The instrument uses a Likert scale and a series of questions that call for either yes or no replies from the respondent.

Validity and Reliability Instruments. University students and faculty members affiliated with the University of Qamarul Huda Badaruddin Bagu were the subjects of the study instrument's testing, conducted on a sample size of 207 individuals. According to the findings of the item validity study, based on the Pearson correlation test with a significance level of $p < 0.05$, it was determined that three items failed to meet the criteria for validity. Consequently, these items were not included in the study that was conducted further. All question items with a validity score of 0.8 or higher were placed in the high validity category, indicating that they are suitable for use as research instruments. This assumption is supported by the reliability analysis performed on all items, as demonstrated by using Cronbach's Alpha and its conclusions. There is a high level of internal consistency and dependability, as indicated by the mean value of the alpha test scale, which is more significant than this value. As a consequence of this, the utilisation of this instrument for the study is highly feasible.

Data Collection. For data collection, an online questionnaire survey was carried out through a Google form, in-depth interviews were performed, and direct practicum observation was carried out.

Data Analysis. Since it is suitable for studying unpaired categorical data, the Kruskal-Wallis statistical analysis method was utilised in this investigation. It evaluated the dependent variable's data on more than two categorical data.

RESULTS AND DISCUSSION

What is the present status of the four universities on science practicum assessment to scientific attitude?

By the information shown in Table 1. The current technique of assessing the science practicum is handled chiefly manually, with most respondents saying that the evaluation is not completed in real-time and lacks feedback, as noted by a considerable proportion of participants (73.70, M=0.24, SD=0.42). In addition, respondents generally agree (71.30 %, mean = 3.04, standard deviation = 0.54) that the practicum evaluation ought to be based on the accomplishment of performance processes. Furthermore, 68.40% of participants (mean=3.18, standard deviation=0.54) indicated that they agreed with the idea that scientific attitude assessment ought to be incorporated into the process of science practice.

Table 1 – Current science practicum assessment at four university

No	Data description	N	%	Mean (M)	SD
1	Practicum Assessment has not been done real-time and without feedback	280	73.70	0.24	0.42
2	Agree if the process of assessment practicum is based on performance assessment	271	71.30	3.04	0.54
3	Agree scientific attitude must be assessed on science practicum	260	68.40	3.18	0.54

The findings are consistent with the information gathered throughout the practice time and the in-depth interviews carried out with the heads of each department associated with the practicum at four different institutions. To be more specific, the findings suggest that no performance evaluation was carried out during the practice, which also included the absence of a scientific attitude evaluation. Furthermore, no attempt was made to provide feedback alongside the assessments. The evaluation procedure had a preliminary test carried out before the commencement of the practice and a final response evaluation carried out after the training was completed. The absence of a standardised evaluation instrument validated and tested for its effectiveness is why assessment is not carried out during practice. This is because there is no such instrument readily available.

Whether the type of university and the respondents' responses have a statistically significant effect on the assessment of scientific attitude ?

There was not a significant difference ($p > 0.05$) in the responses of teachers and students about the assessments of scientific attitude, as evidenced by the statistical analysis that was carried out using the Kruskal-Wallis test (Table 3). It can be deduced from this that both groups agree that the bulk of scientific attitude assessments were not carried out during the science practicum. The reason for this is that there were no instruments that were suitable and trustworthy for measuring scientific attitude throughout the entirety of the process of the investigation.

Table 2 – Scientific attitude measurement based on respondent and university variables

Variables	Category	N	Mean (M)	SD	Sig. Kruskal – Wallis
Respondent	Never	263	1.72	0.45	0.15
	Seldom	100	1.69	0.46	
	Often	13	1.25	0.45	
University	Never	265	5.41	0.49	0.13
	Seldom	99	5.57	0.49	
	Often	12	5.50	0.55	

This finding was supported by the findings of in-depth interviews that were carried out with the management level of each department that was associated with academic and practicum. During these interviews, it was discovered that the individuals indicated above only possessed non-standardised instruments for evaluating the practicum opportunity.

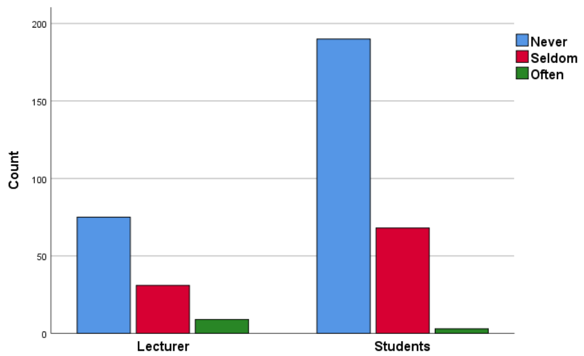

Figure 1 – Response Lecturer and Student on Scientific attitude Measurement

Table 3 – Kruskal-Wallis statistical analysis on Scientific Attitude Measurement of Respondent Responses

Total N	376
Test Statistic	7.007
Degree Of Freedom	2
Asymptotic Sig. (2-sided test)	.130

The evidence shown in Figure 2 indicates that the four universities did not carry out scientific attitude assessments while putting the science practicum into action. This finding is further corroborated by the findings of a statistical test (Table 3), which reveals that there is no statistically significant difference ($p > 0.05$) between the kind of university and the measurement of scientific attitude. This implies that there is no correlation between the two characteristics. The majority of the scientific attitude measurements (Figure 1) had not been carried out at the time when the practice was being conducted. As a consequence of the absence of assessment instruments throughout the practice phase, the reliance on practicum observations has become necessary. In order to make up for this constraint, in-depth interviews were carried out with informants employed by four different educational institutions. The significance of scientific attitude assessments was brought to light by these interviews, yet it was observed that there was a shortage of standardised and validated instruments for validity and reliability. Consequently, the lack of such instruments creates difficulties in establishing accountability for the results obtained.

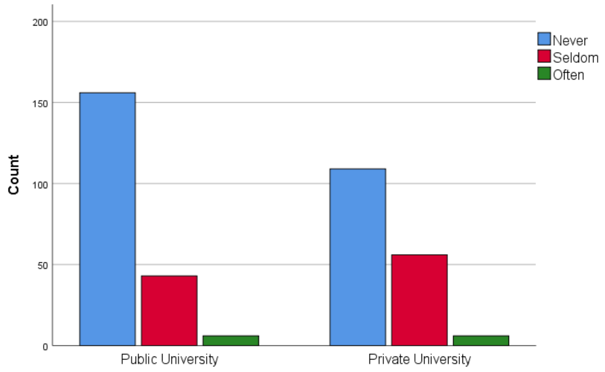

Figure 2 – Response University type on Scientific Attitude Measurement

Table 4 – Kruskal-Wallis statistical analysis for Scientific Attitude Measurement by university variables

Total N	376
Test Statistic	11.794
Degree Of Freedom	2
Asymptotic Sig. (2-sided test)	.150

Using questionnaires, in-depth interviews, and observations demonstrated that a severe constraint hampers the evaluation of the scientific practicum. This limitation was identified through the utilisation of these methods. The absence of a complete performance evaluation is a limitation linked to the absence of measurement instruments that contain both validity and reliability. This limitation pertains to the lack of such an assessment. When it comes to the Android platform, it is recommended that access be restricted to just authorised workers to supervise the practicum assessment method.

The Current Science Practicum Assessment. In-depth interviews were carried out at four different institutions at the management level. Academics deans and laboratories directors were the subject for these interviews. The results of these interviews are in Table 1. To verify results that fill out data from a questionnaire, in-depth interviews are conducted. Findings from in-depth interviews revealed that there was no difference in the practicum assessment methods among the four colleges. These universities included both public and private institutions.

The initial response (pre-test), the final response (post-test), and the final report were all evaluated by teaching assistants as part of the practicum assessment process. The evaluation results are first reviewed and approved by the head of the laboratory, and then they are handed over to the course instructor. Based on the observations, it has been determined that the evaluation is still carried out manually, and no performance evaluation instrument has been discovered. As a result, it is essential to design performance evaluation tools that are both valid and trustworthy to evaluate the process of putting the practicum into action. These results are in line with what was found during the detailed interview with the lab director, and the academic dean's. An assessment of performance has not been conducted, which is deemed to have limited engagement, cooperation without the use of legitimate standardized tools, and reliability.

It has been determined from interviews conducted at four different tertiary institutions that a performance assessment procedure has not supported the implementation of the practicum. As a result, the practicum implementation has produced less effective results [1, 10, 6] respectively. Most evaluations are carried out through reports one week after the conclusion of the practicum. This is because skills are essential to activities that are not monitored or quantified during the internship. As a result, we require performance evaluation tools that can be accessed digitally and evaluated in a short amount of time. These tools should be legitimate, consistent, and dependable to reduce the amount of work, time, and money required [17]. Because lecturer assistants tend to rely on their memories when administering evaluations, it is advised that a performance evaluation tool be developed during the practicum year [3].

It is stated in [6] the majority of lecturers assess the student's practicum process following the completion of the practicum's activities. Measuring should begin at the start of the practicum and continue until it is complete; evaluation should not be done after the practicum activities have concluded. This means that at each level of the practicum, a thorough evaluation should be carried out, keeping an eye on all areas, particularly the cognitive, psychomotor, and affective [1, 21]. In addition, [6] stated that performance eval- uation is not only carried out to evaluate students' cognitive capabilities, but it is also necessary to evaluate their abilities and attitudes. To carry out this measurement, it is essential to have a performance evaluation instrument that is not only valid but also dependable and easy to comprehend.

Real-time Assessment Scientific Attitude. The study's outcomes reveal that both lecturers and students have preferred practitioners to engage in real-time practices using digital technology, notably through the accessibility of mobile phone applications and Android and web applications. The driving force behind this preference is the need to improve access to scientific attitude evaluations and the capability to obtain results and feedback anytime and anywhere without any restrictions. This is consistent with the statement made in [16], which states that a mobile application enables real-time examination of assessment outcomes across various levels and formats by the instructor's preferences, which may include numerical, descriptive, or dual approaches. Furthermore, the application is built specifically for evaluating scientific attitudes and generates descriptive assessment reports to assess students' performance. The application provides educators with data that is both comprehensive and systematic, which equips them with the ability to clarify instructional and learning goals, evaluate student grades based on explicit criteria and competencies, accurately summarise student performance, and identify patterns of proficiency and deficiencies in student work. It was to utilise the descriptive Rubric Score application inside the educational context, namely in the arena of evaluating student success by teachers working in elementary and secondary schools, that the application was developed. The tool performs real-time analysis of assessment results at multiple levels, which are determined by the teacher (for example, student, department, class, and school). Depending on the instructor's preferences, the results are presented in either a numerical or descriptive manner or a combination of the two. Reading comprehension, writing, scientific attitude, participation-collaboration, perseverance, and computational thinking are the six characteristics evaluated using the Stratified Criteria Scale. This scale defines evaluative criteria and evaluates students based on these characteristics.

The ability to conveniently transfer devices, the responsiveness of operating systems and applications, the capability to link users across multiple time zones and locales, and the facilitation of social contact are just some of the many benefits consumers can enjoy due to mobile technology. The rapid adoption of mobile applications and technological advancements in the field of education is a phenomenon that is occurring all around the world [5]. The academic community, and notably those working in education, find mobile technology quite appealing [8]. The term "mobile learning" refers to utilising mobile technology, including computers, laptops, mobile phones, audio players, and electronic books, to achieve the goal of electronic learning [1]. By utilising internet-based platforms and technological breakthroughs, mobile learning makes it easier for students to work together on their education and fosters the sharing of ideas. This allows students to overcome traditional classroom settings' time and location limitations [4]. Students can now access course materials and engage with educational content dynamically and interactively because of the widespread availability of mobile devices that are always connected to the internet.

In the twenty-first century, educators' most critical challenge is maintaining student interest and participation in effective learning through mobile devices [22]. The significance of fostering advanced cognitive qualities in pupils, such as problem-solving and a scientific mindset, has been widely acknowledged [7]. Mobile learning is one of the most cutting-edge educational options that has recently arisen [18]. Learning through mobile devices has been found to influence both students and teachers positively. This is because mobile learning makes it easier for students to comprehend the content being taught and improves various cognitive capacities, including communication, problem-solving, creativity, and higher-order thinking capabilities.

A set of tools that provide a variety of answers to problems encountered in education is one definition of technology that is sometimes used [9]. "Education is intended to cultivate individuals who possess a wide range of knowledge, demonstrate creativity, demonstrate proficiency in the utilisation of digital technologies, and possess the ability to adapt to changing circumstances" [18] educational institutions strive to accomplish. Furthermore, the exploitation of information technology in academic settings, including the internet and multimedia systems, has been meant to improve the quality of learning by making it easier for students to access vital materials and develop a scientific mindset [19].

Three distinct categories are included in the examination of higher-order thinking skills (HOTS): the ability to apply concepts to different contexts, the classification of higher-order thinking skills as a type of knowledge, and the ability to establish connections with other people in unexpected conditions. Additionally, developing a scientific mindset comprises the ability to comprehend logical intricacies, engage in reflective thinking, and engage in argumentation about generating informed judgments or conclusions concerning the matter at hand. In addition, one of the most critical aspects of an individual's talents is their problem-solving ability, especially their capacity to create creative, unusual, and unique [20].

Performance Assessment and Feedback. According to the statistical analysis findings, there was no statistically significant connection between the assessments that teachers carried out and the student's performance. When the outcomes of the statistical evaluation of academic achievement at both public and private institutions were compared, similar conclusions were discovered in both types of colleges. It was decided to conduct the performance evaluation without providing any comments to the individual. A combination of in-depth interviews with management staff and practicum helpers and first-hand observations of the practicum provided evidence to support this assertion. The utilisation of a digital application is considered to be the most advantageous when it comes to enhancing performance evaluation and providing automatic feedback that both students and teachers can easily access. According to [23, 22], the automated program generates Descriptive Assessment Reports to evaluate student performance. Researchers utilise these reports to assess student proficiency in scientific attitude. The researchers introduced an additional variable into the analytics rubric that was pre-installed on the program. Formative guidance provides comments and recommendations on improving the work to satisfy the necessary standards [22, 7, 11, 9]. Assessment is the process of evaluating students' performance in relation to a certain aim, and formative guidance is the feedback provided. The term "feedback" refers to the difference between the actual values of the system's parameters and the reference values of those parameters. This difference is then used to adjust the discrepancy in a particular way.

The study began with the intention of assisting educators in improving their instructional practices and students' learning outcomes through implementing formative and descriptive assessment strategies. This approach effectively communicates the expectations placed on students about their performance. Secondly, it offers pupils helpful feedback, which assists them in gaining an awareness of their areas of strength and areas they may develop for themselves. Furthermore, it guarantees that the evaluation process is carried out uniformly and impartially, thereby fostering equity among the pupils. Last but not least, it has been demonstrated that the utilisation of rubrics results in improved student learning and encourages self-evaluation [12]. In addition, the use of rubrics makes it easier for teachers to accomplish the following tasks: a) defining clear objectives for teaching and learning; b) evaluating student grades based on specific criteria and skills; c) consistently articulating student performance; and d) recognising patterns of strengths and weaknesses in student work [20]. Employing analytical rubrics that are suitable for the subject matter at hand may potentially improve the dependability of performance evaluations.

CONCLUSIONS

Standardised instruments for assessing the scientific attitude with real-time feedback at science practicum are unavailable. Therefore, it is necessary to develop valid and reliable performance assessment instruments to determine the implementation process that provides real-time feedback on science practicum.

REFERENCES

1. Bensley, D. A., Masciocchi, C. M., & Rowan, K. A. (2021). A comprehensive assessment of explicit critical thinking instruction on recognition of thinking errors and psychological misconceptions. *Scholarship of Teaching and Learning in Psychology, 7*(2), 107–122. doi: 10.1037/stl0000188

2. Chafiq, N., Talbi, M., & Ghazouani, M. (2018). Design and Implementation of a Risk Management Tool: A Case Study of the Moodle Platform. *International Journal of Advanced Computer Science and Applications, 9*(8). doi: 10.14569/ijacsa.2018.090858

3. Dhina, M. A., Hadisoebroto, G., Mubaroq, S. R., & Gustiana, I. (2021). Implementation of digital performance assessment to measure pharmacy physics laboratory skills. *Momentum: Physics Education Journal*, 65–72. doi: 10.21067/mpej.v5i1.5146

4. Harsch, C., Seyferth, S., & Villa Larenas, S. (2021). Evaluating a collaborative and responsive project to develop language assessment literacy. *Language Learning in Higher Education, 11*(2), 311–342. doi: 10.1515/cercles-2021-2020

5. Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1), 81–112. doi: 10.3102/003465430298487

6. Kruit, P. M., Oostdam, R. J., van den Berg, E., & Schuitema, J. A. (2018). Assessing students' ability in performing scientific inquiry: instruments for measuring science skills in primary education. *Research in Science & Technological Education*, 1–27. doi: 10.1080/02635143.2017.1421530

7. Li, X., & Liu, Q. (2020). Social Media Use, eHealth Literacy, Disease Knowledge, and Preventive Behaviors in the COVID-19 Pandemic: Cross-Sectional Study on Chinese Netizens. *Journal of Medical Internet Research, 22*(10), e19684. doi: 10.2196/19684

8. Liu, H., Liu, Z., Lai, Y., & Li, L. (2021). Factors Influencing Collaborative Innovation Project Performance: The Case of China. *Sustainability, 13*(13), 7380. doi: 10.3390/su13137380

9. Montenegro-Rueda, M., Luque-de la Rosa, A., Sarasola Sánchez-Serrano, J. L., & Fernández-Cerero, J. (2021). Assessment in Higher Education during the COVID-19 Pandemic: A Systematic Review. *Sustainability, 13*(19), 10509. doi: 10.3390/su131910509

10. Riantini, N. L. R., Suastra, I. W., & Budi Adnyana, P. (2018). Development of science practicum performance assessment in junior high school. *SHS Web of Conferences, 42*, 00090. doi: 10.1051/shsconf/20184200090

11. Ronau, R. N., Rakes, C. R., & Niess, M. L. (Eds.). (2012). *Educational Technology, Teacher Knowledge, and Classroom Impact*. doi: 10.4018/978-1-60960-750-0

12. Saleh, M., & Mohamed, R. (2018). Recommendations for Building Adaptive Cognition-based E-Learning. *International Journal of Advanced Computer Science and Applications, 9*(8). doi: 10.14569/ijacsa.2018.090850

13. Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students in a blended medical education course. *Medical Teacher, 39*(7), 757–767. doi: 10.1080/0142159x.2017.1309376

14. Sudirman, S. (2020). Pengembangan Aplikasi Untuk Pemantauan Dan Evaluasi Pendidikan Anak Usia Dini [Application Development for Monitoring and Evaluation of Early Childhood Education]. *INFOTECH: Jurnal Informatika & Teknologi, 1*(2), 107–115. doi: 10.37373/infotech.v1i2.61 (in Indonesian).

15. Sudirman, S., Sarjan, M., Rokhmat, J., Hamidi, H., Muliadi, A., Azizi, A., Fauzi, I., Yamin, M., Muttaqin, Muh. Z. H., Rasyidi, M., Ardiansyah, B., Khery, Y., & Rahmatiah, R. (2022). Praktik Penilaian Guru Pendidikan Sains antara Keyakinan atau Pengetahuan Guru? Perspektif Filsafat. *Jurnal Ilmiah Profesi Pendidikan, 7*(3c), 2018–2025. doi: 10.29303/jipp.v7i3c.889

16. Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2017). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education, 76*(3), 467–481. doi: 10.1007/s10734-017-0220-3

17. Tseng, J.-J. (2014). Developing an instrument for assessing technological pedagogical content knowledge as perceived by EFL students. *Computer Assisted Language Learning, 29*(2), 302–315. doi: 10.1080/09588221.2014.941369

18. Walters, S., Silva, P., & Nikolai, J. (2017). Teaching, Learning, and Assessment: Insights into Students' Motivation to Learn. *The Qualitative Report*. doi: 10.46743/2160-3715/2017.2777

19. Wiethe-Körprich, M., & Bley, S. (2017). Prospective educators as consumers of empirical research: an authentic assessment approach to make their competencies visible. *Empirical Research in Vocational Education and Training, 9*(1). doi: 10.1186/s40461-017-0052-5

20. Yan, E. M. Y. (2022). Using Online Peer Assessment Activities to Enhance Team Collaboration in Two Undergraduate Courses. *Innovative Teaching and Learning, 4*(1), 59–73. doi: 10.4208/itl.20220104

21. Yan, Z. (2019). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & Evaluation in Higher Education, 45*(2), 224–238. doi: 10.1080/02602938.2019.1629390

22. Zhang, L., & Li, P. (2022). Problem-Based mHealth Literacy Scale (PB-mHLS): Development and Validation. *JMIR MHealth and UHealth, 10*(4), e31459. doi: 10.2196/31459

23. Zhao, D., Chis, A., Muntean, G.-M., & Muntean, C. H. (2018). A large-scale pilot study on game-based learning and blended learning methodologies in undergraduate programming courses. *Edulearn18 Proceedings*. doi: 10.21125/edulearn.2018.0948