

# The Development of Performance Assessment Inventory (PAI) to Access Students' Self and Peer Assessment in the Innovation Course

Puri Selfi Cholifah <sup>1,\*</sup>, Ni Luh Sakinah Nuraini <sup>1</sup>, Putri Mahanani <sup>1</sup>

<sup>1</sup> Dept. of Primary and Preschool Teacher Education, Faculty of Education, Universitas Negeri Malang, Malang, Indonesia

\*Corresponding author. Email: [puri.selfi.fip@um.ac.id](mailto:puri.selfi.fip@um.ac.id)

## ABSTRACT

This research aims to develop an instrument performance assessment rubric as a first step for further development. This research procedure adapts the ADDIE development model, a development model with five stages: analysis, design, development, implementation, and evaluation. The R & D stages are described as follows. The subjects involved in the development of this product were two lecturers as assessment experts and learning experts and 14 elementary school teacher education students of the State University of Malang who had taken the Educational Innovation Diffusion course. Data collection was carried out from August to November 2020. The data collection methods used were observation, questionnaires, documentation, and tests. Based on the result of the research, the average score of about 90% was obtained for the assessment and learning experts. The instrument developed was in the category of feasible use. Moreover, further data from the small group test shows that the instrument designed can be used to clarify the task and the rubric's readability with a final score of around 94%.

**Keywords:** performance assessment inventory, self-assessment, peer-assessment, innovation course

## 1. INTRODUCTION

Assessment activities are a crucial part of the learning process. The urgency of alternative assessment is becoming a trend among higher education institutions, where the traditional assessment (pencil and paper test) is deemed insufficient to access higher-order thinking skills [1], [2]. Moreover, in higher education, students are required to show their performance to achieve higher education output in the world of work.

However, in the implementation of ongoing lectures, access to assess performance independently is still hampered due to the absence of an appropriate instrument to assess student performance. This is supported by data that during the period August to December 2019, where odd semester lectures were held, observational studies were carried out related to students' ability to assess their performance and their peers' performance, but the results were not optimal. Based on data from the initial instrument developed, it turns out that students have not been able to independently carry out the performance appraisal. Moreover, in its implementation, when the assessment is carried out in manual form (using paper), the results of the interviews further show weaknesses.

One of the most important is that students who are not present cannot provide peer assessments to their friends, so the data obtained is for students who appear incomplete.

In this regard, one of the efforts that can be made to contribute to student access rights in assessing individual and peer performance is developing a performance assessment inventory (PAI). With this instrument, students will find it easier to access self and peer assessments, especially in innovation courses. In the adult learner construct, one of the critical factors to support student learning and involvement is self-assessment and peer assessment [3] - [5]. However, in the Innovation course (either in the form of Diffusion of Innovation or Innovation Management Course), peer contributions and objective assessments are needed regarding the nature of oneself and peers' innovativeness.

The line of thought above forms the basis that it is necessary and has never developed an instrument related to performance assessment inventory. The urgency of this development and answering the challenges of the need for innovation courses are also to answer challenges related to independent ability in peer and self-assessment

as part of the 21st Century learning development, which emphasizes higher-order thinking skills. Besides, with this instrument, it is hoped that it can become a forum for students to choose their learning preferences, especially in carrying out a digital-based assessment process that is more fun and following 21st-century students' characteristics. Therefore, this research was carried out. This research aims to develop an instrument performance assessment rubric as a first step for further development.

## **2. LITERATURE REVIEW**

### **2.1. Performance Assessment Inventory (PAI)**

The current era emphasizes a change in the assessment paradigm from what was previously only done traditionally to an alternative assessment. In traditional assessments, more emphasis is placed on pencil and paper-based assessments or cognitive skills. Although some felt that the conventional assessment selection was considered more accurate [6]. The study also shows that it is quite common in some university systems to use this assessment, although there are several drawbacks, including high cost and time-consuming process [7].

In performance appraisal, the main focus is higher-order thinking skills that integrate the understanding of various subjects [2]. This is because performance-based learning skills focus not only on learning but also on independent learning [8]. Basically, there has been an increase in tertiary institutions that require further assessment, namely the implementation of performance appraisals. However, some educators feel heavy because the workload is more than just assessing written test assignments [9]. Nonetheless, it appears that students view performance appraisals as reinforcing comprehension and are mainly seen as more authentic and reactive in terms of their learning needs [1].

### **2.2. Self-Assessment**

The current era emphasizes a change in the assessment paradigm from what was previously only done traditionally to an alternative assessment. In traditional assessments, more emphasis is placed on pencil and paper-based assessments or cognitive skills. Although some felt that the conventional assessment selection was considered more accurate [6]. The study also shows that it is quite common in some university systems to use this assessment, although there are several drawbacks, including high cost and time-consuming process [7]. In performance appraisal, the

### **2.3. Peer Assessment**

Peer assessment or peer assessment is generally an assessment process in which students assess peers' work [12]. The study shows several positive impacts of using

peer assessment, especially at the tertiary level [12], [13]. Further studies show that students feel that they can benefit from the inspiration of the work they do, immensely to improve the work they do [4], [5].

## **3. METHOD**

This research procedure adapts the ADDIE development model, a development model with five stages: analysis, design, development, implementation, and evaluation. The R & D stages are described as follows.

First, the analysis stage. In this stage, the student needs analysis includes the level of innovation initiation and the innovation initiation instrument [14]. This analysis consists of the student's target performance assessment inventory (PAI). Also, in this analysis stage, data related to the field's situation were extracted by exploring data related to existing digital assessment instruments. The next analysis carried out is related to the study of the rubric component and the assessment task contained in this product.

Second, the design stage. In this stage, the design includes making the instrument layout design, assigning tasks and rubrics. Next, is the development stage. At the development stage, several components will be a significant part of product development, which includes: (1) product development, (2) expert validation, (3) validation by practitioners. In the development stage, inventory performance assessment. Regarding expert validation, this product is validated by learning experts and assessment experts. Validation of practitioners, among others, is a lecturer in the Innovation Management MK who will participate in the development of this product by providing assessments and suggestions.

Fourth, the implementation stage. The implementation stage consists of a small-scale product trial stage. In this trial, users will see the instrument's practicality consisting of the PAI assignment and rubric. In this case, 14 students were involved in small-scale trials.

Lastly, the evaluation stage. The evaluation stage consists of formative and summative evaluations. Formative evaluation is carried out at all stages of development, from analysis to implementation. The summative evaluation stage is an activity to review the achievements of product development carried out, and one of them is through the experimental analysis activities carried out. Still, this study will only reach the formative evaluation stage.

The subjects involved in the development of this product were two lecturers as assessment experts and learning experts and 14 elementary school teacher education students of the State University of Malang who had taken the Educational Innovation Diffusion course. The subject selection was based on purposive sampling

technique after exploring data related to students' conditions and initial characteristics who took the Educational Innovation Diffusion Court. Data collection was carried out from August to November 2020. The data collection methods used were observation, questionnaires, documentation, and tests.

The data were obtained from the product development process data, which included assessments and suggestions from experts, practitioners and users, which were analyzed descriptively by presenting percentages and recommendations in qualitative data. The following data is data related to the score from the performance assessment inventory (PAI) instrument tested for validity.

#### 4. RESULT

At the development stage, the performance assessment inventory (PAI) instrument was developed by prioritizing the education innovation course's characteristics. This instrument consists of two main parts, namely the performance task and the self-assessment rubric. In the performance task section, the main components seen are Title; Subject; Level; Time; Instructions; Output target; Skills required; Materials needed; and assessment.

The performance task section in each task (task 1, task 2 and task 3) has a different description. In task 1 (analysis of innovation needs, this project becomes a starting point for the next innovation task. This task is a process of extracting data (interviews, observation, and documentation) from the needs for innovation in the field (innovation targets). Furthermore, in task 2 (innovation development), this project is intended to develop innovation from the basic step. The students are expected to determine: (1) initial condition analysis, (2) type of innovation, (3) innovation attributes, (4) innovation development, (5) innovation instruments, (6) innovation barriers and support, (7) innovation diffusion targets, (8) innovation diffusion activities, (9) innovation diffusion budget In Task 3 (innovation posters), this task puts forward the dissemination of products through making posters so that innovative products are designed to be disseminated.

All assignments are accompanied by an assessment rubric consisting of a scale and dimensions of the assessment. In the scale component, this rubric uses a

scale of 4 scales, namely 4 (extraordinary), 3 (good), 2 (enough), and 1 (below standard). The dimensions in rubric one is related to ideas, data, and writing. The second assignment, rubric includes idea, organization, and writing. While task 3 is related to the poster, it includes content, organization, and appearance.

After the product is developed, in the formative evaluation stage, validation is carried out with an assessment expert validator and a learning expert. In the expert section of the assessment, the assessed components are related to five main aspects: compliance with the curriculum, clarity of performance tasks, opportunities for the assessment process, suitability of performance rubrics, and accuracy in using language. As for the learning validation section, the assessed components include conformity to the curriculum, suitability for learning, suitability for exposure to innovation, and accuracy in using language.

After improving the rubric and performance tasks on the product being developed, it is continued to a small-scale test to assess the rubric's usability (the readability of the assessment instrument). At this stage, the small group test involved 14 elementary school teacher education students who had taken the educational innovation diffusion course. The data on this assessment include clarity of performance tasks and rubric readability. The data related to the expert validation assessment results and the results of small group trials are presented in Table 1.

Based on the results of expert validation, it was stated that the average value of the assessment instrument product was 90.0%, with the highest component in the task of developing innovation. The learning expert as the validator gave a score of 90.7% with the highest score criteria. The same as the assessment expert, namely in task component 2 (innovation development), received 94.4%.

Based on these data, it can be stated that the assessment instrument product developed is in the valid category, so it is suitable for use. In the small group test, data related to users were obtained, different things were seen in the highest value of the product that fell to task 3, namely the innovation poster. However, the average product developed obtained a score of 94.1%, so it can be stated that the instrument can be used based on its legibility element.

**Table 1 Recapitulation of The Results of The Validation and Small-Scale Group Test**

No	Instrument	Assessment Expert	Learning Expert	User (Small Scale Group)
1	TASK 1 (analysis of innovation needs)	88.8%	88.9%	92.5%
2	TASK 2 (Innovation development)	91.3%	94.4%	93.8%
3	TASK 3 (Poster)	90.0%	88.9%	96.0%
	<b>Average</b>	<b>90.0%</b>	<b>90.7%</b>	<b>94.1%</b>

Qualitative data are also obtained in the process of development and formative evaluation. Data is received in the comment's column for each assessment. For example, related to the instrument's readability element, related suggestions for the use of language are more concise or not long-winded. All qualitative data are then used as material for comprehensive manuscript improvement to obtain an adjusted final product.

## 5. DISCUSSION

Based on expert investigations, it is obtained that the developed instrument is in a suitable category. The use of performance assessment within the university's scope does require more effort, especially in terms of time. In line with this, the performance task of the innovation development process carried out for 12 weeks shows that the performance task lasts throughout the lecture. This is also in line with studies that state that universities' performance appraisal system is time-consuming [7], [15].

In performance appraisal, the main focus is higher-order thinking skills that integrate the understanding of various subjects. This is as presented in the product of innovation development, including the skills needed. For example, in section 1, the skills needed to complete these tasks include synthesizing information, collecting data, conducting research, writing research results. This is in line with studies showing that performance appraisal is closely related to higher-order thinking skills [2], [16], [17]. This is because performance-based learning skills focus not only on learning but also on independent learning [8].

Regarding the data from small-scale trials, it was found that students' enthusiasm in the process of assessing the performance of the investment seemed high. This was indicated by the data related to the clarity of the innovation task and the rubric's suitability performance. This is in line with the opinion that students see performance appraisal as reinforcing understanding and are specifically felt to be more authentic and reactive concerning their learning needs [1].

In practice, assigned assignments lead to harnessing 21st-century skills that are needed today. What's more, tasks that are linked to the needs of solving surrounding problems are more authentic than using only conceptual tasks.

In the qualitative data obtained, some suggest shortening the rubric so that it is not too long. This is because, in practice, scoring using a rubric requires its accuracy. Nonetheless, the detailed assessment rubric criteria make a better interpretation of the performance performed [18]. Moreover, the rubric-making system that can be done together with teachers and students makes performance better and fairer. This is an essential thing in the assessment procedure carried out [19]. It also

makes the performance task clearer with the criteria / levels discussed earlier [20].

## 6. CONCLUSION

Considering the data from the expert validity test results where an average score of about 90% was obtained for the assessment and learning experts, the instrument developed was in the category of feasible use. Moreover, further data from the small group test shows that the instrument designed can be used to clarify the task and the rubric's readability with a final score of around 94%. However, this stage is the initial stage of developing a performance assessment inventory, which will continue to develop a digital rubric that uses a mobile basis to facilitate self-assessment and peer assessments in education innovation courses.

## REFERENCES

- [1] P. Iannone dan A. Simpson, "Students' views of oral performance assessment in mathematics: straddling the 'assessment of' and 'assessment for' learning divide," *Assessment & Evaluation in Higher Education*, vol. 40, no. 7, hlm. 971–987, Okt 2015, doi: 10.1080/02602938.2014.961124.
- [2] B. S. Potter, J. V. Ernst, dan E. J. Glennie, "performance-based assessment in the secondary STEM classroom," *Technology and Engineering Teacher*, vol. 76, no. 6, hlm. 18, 2017.
- [3] M. Taras, "Student self-assessment: processes and consequences," *Teaching in Higher Education*, vol. 15, no. 2, hlm. 199–209, Apr 2010, doi: 10.1080/13562511003620027.
- [4] T. Wanner dan E. Palmer, "Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback," *Assessment & Evaluation in Higher Education*, vol. 43, no. 7, hlm. 1032–1047, Okt 2018, doi: 10.1080/02602938.2018.1427698.
- [5] D. Weaver dan A. Esposto, "Peer assessment as a method of improving student engagement," *Assessment & Evaluation in Higher Education*, vol. 37, no. 7, hlm. 805–816, Nov 2012, doi: 10.1080/02602938.2011.576309.
- [6] C. J. Dommeyer, P. Baum, K. S. Chapman, dan R. W. Hanna, "Attitudes of Business Faculty Towards Two Methods of Collecting Teaching Evaluations: Paper vs. Online," *Assessment & Evaluation in Higher Education*, vol. 27, no. 5, hlm. 455–462, Sep 2002, doi: 10.1080/0260293022000009320.
- [7] M. Lalla dan D. Ferrari, "Web-based versus paper-based data collection for the evaluation of teaching activity: empirical evidence from a case study," *Assessment & Evaluation in Higher Education*, vol. 36, no. 3, hlm. 347–365, Mei 2011, doi: 10.1080/02602930903428692.
- [8] C. Stephens-Himonides, M. Hilley, dan M. Hilley, "Beyond the Classroom," *The Routledge Companion to*

- Music, Technology, and Education, Jan 20, 2017. <https://www.taylorfrancis.com/> (diakses Jan 13, 2020).
- [9] R. Cummings, C. D. Maddux, dan A. Richmond, "Curriculum-embedded performance assessment in higher education: maximum efficiency and minimum disruption," *Assessment & Evaluation in Higher Education*, vol. 33, no. 6, hlm. 599–605, Des 2008, doi: 10.1080/02602930701773067.
- [10] H. Andrade dan A. Valtcheva, "Promoting Learning and Achievement Through Self-Assessment," *Theory Into Practice*, vol. 48, no. 1, hlm. 12–19, Jan 2009, doi: 10.1080/00405840802577544.
- [11] L. Leach, "Optional self-assessment: some tensions and dilemmas," *Assessment & Evaluation in Higher Education*, vol. 37, no. 2, hlm. 137–147, Mar 2012, doi: 10.1080/02602938.2010.515013.
- [12] N. Falchikov dan J. Goldfinch, "Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks," *Review of Educational Research*, vol. 70, no. 3, hlm. 287–322, Sep 2000, doi: 10.3102/00346543070003287.
- [13] D. M. A. Sluijsmans, S. Brand-Gruwel, dan J. J. G. van Merriënboer, "Peer Assessment Training in Teacher Education: Effects on performance and perceptions," *Assessment & Evaluation in Higher Education*, vol. 27, no. 5, hlm. 443–454, Sep 2002, doi: 10.1080/0260293022000009311.
- [14] E. Chell dan R. Athayde, *The identification and measurement of innovative characteristics of young people: Development of the youth innovation skills measurement tool*. NESTA, 2009.
- [15] M. B. Kane dan R. Mitchell, *Implementing performance assessment: Promises, problems, and challenges*. Routledge, 2013.
- [16] N. R. Council, *Assessing 21st century skills: Summary of a workshop*. National Academies Press, 2011.
- [17] E. Van Laar, A. J. Van Deursen, J. A. Van Dijk, dan J. De Haan, "The relation between 21st-century skills and digital skills: A systematic literature review," *Computers in human behavior*, vol. 72, hlm. 577–588, 2017.
- [18] J. McCarthy, "Evaluating written, audio and video feedback in higher education summative assessment tasks," *Issues in Educational Research*, vol. 25, no. 2, hlm. 153–169, 2015.
- [19] V. Ulker, "The Design and Use of Speaking Assessment Rubrics," *Journal of Education and Practice*, vol. 8, no. 32, 2017.
- [20] D. Balch, R. Blanck, dan D. H. Balch, "Rubrics–Sharing the Rules of the Game.," *Journal of Instructional Research*, vol. 5, hlm. 19–49, 2016.