MULTILATERAL: Jurnal Pendidikan Jasmani dan Olahraga
23 (1), 2024, 1-14
Available Online: https://ppjp.ulm.ac.id/journal/index.php/multilateralpjkr

Development of performance evaluation instruments for faculty of national sports sciences lecturers

Pengembangan instrumen evaluasi kinerja dosen fakultas ilmu keolahragaan nasional

Masrun[1], Umar[2], Khairuddin[3]
[1,2,3] Fakultas Ilmu Keolahragaan, Universitas Negeri Padang, Indonesia
Email: masrun@fik.unp.co.id[1], umarkepel@fik.unp.ac.id[2], khairuddins2@fik.unp.ac.id[3]

ABSTRACT

Assessing the performance of lecturers is a vital tool to ensure optimal education and learning aligned with established quality standards. However, there is currently no universally accepted tool for evaluating the performance of Faculty of Sport Science (FIK) lecturers in Indonesia. This research aims to create and standardize a performance evaluation instrument for FIK lecturers, focusing on Professional, Social, and Personal aspects. Employing a Research and Development (R&D) design with a procedural model, the standardization process involves several key steps. The study focuses on UNP FIK lecturers as test subjects. The standardization process encompasses compiling a framework, expert testing, a limited trial involving 30 samples, and a broader test with 82 samples. Practicality and effectiveness tests are also conducted. Data analysis, performed using the IBM SPSS program with a significance level of $p < 0.05$, supports the formation of a standardized instrument capable of nationally evaluating the performance of FIK lecturers.

***Key words**: evaluation instrument; lecturer performance; sport science.*

Menilai kinerja dosen adalah alat yang penting untuk memastikan pendidikan dan pembelajaran yang optimal sesuai dengan standar kualitas yang telah ditetapkan. Namun, saat ini belum ada alat yang dapat diterima secara universal untuk mengevaluasi kinerja dosen Fakultas Ilmu Keolahragaan (FIK) di Indonesia. Penelitian ini bertujuan untuk membuat dan menstandarkan instrumen evaluasi kinerja dosen FIK, dengan fokus pada aspek Profesional, Sosial, dan Personal. Menggunakan desain Research and Development (R&D) dengan model prosedural, proses standarisasi melibatkan beberapa langkah utama. Penelitian ini berfokus pada dosen FIK UNP sebagai subjek uji coba. Proses standarisasi meliputi penyusunan kerangka kerja, uji ahli, uji coba terbatas yang melibatkan 30 sampel, dan uji coba yang lebih luas dengan 82 sampel. Uji kepraktisan dan keefektifan juga dilakukan. Analisis data yang dilakukan dengan menggunakan program IBM SPSS dengan taraf signifikansi $p < 0,05$ mendukung terbentuknya instrumen terstandar yang mampu mengevaluasi kinerja dosen FIK secara nasional.

Kata kunci: instrumen evaluasi; kinerja dosen; ilmu keolahragaan.

INFO ARTIKEL

Riwayat Artikel:

Diterima : 25 Oktober 2023
Disetujui : 2 Februari 2024
Tersedia secara *online* Februari 2024
Doi: http://dx.doi.org/10.20527/multilateral.v23i1.17705

Alamat Korespondesi:

Masrun
Jl. Prof. Dr. Hamka, Air Tawar Bar., Kec. Padang Utara, Kota Padang, Sumatera Barat 25171
Email: masrun@fik.unp.co.id

INTRODUCTION

Higher education, particularly within the National Faculty of Sport Sciences, plays a pivotal role in molding and enhancing human resources in the sports field (Aziz et al., 2023a; Aziz et al., 2023b). In supporting the efficacy of the educational process, the role of lecturers, serving as the primary facilitators, is

of utmost importance (Bangun et al., 2023: Wulandari et al., 2023). Evaluating lecturer performance becomes a critical tool to ensure that education and learning occur optimally, adhering to established quality standards. This is underscored by the assertion of Ramli and Jalinus (2013) that the presence of quality educators is an essential prerequisite for the existence of quality education systems and practices.

In the realm of higher education, the primary objective is to facilitate learning between students and educators (Nygaard & Belluigi, 2011). As per Widaningrum (2016), the Instructor/Teacher Performance Evaluation System serves as an evidence-based evaluation mechanism to assess the performance level of each instructor/teacher in fulfilling their primary function as professional educators. Evaluating the work performance of lecturers/teachers should carry a positive significance, aiming to enhance and elevate the professionalism of educators, with a consequential impact on improving student academic achievement (Mahmud, 2018; Mashud et al., 2021).

In the setting of the National Faculty of Sport Sciences, there exists a necessity for a pertinent and all-encompassing lecturer performance evaluation tool to offer an accurate portrayal of lecturers' contributions toward attaining the educational goals and mission of the faculty. Lecturer performance evaluation involves a formal and structured system aimed at gauging, appraising, and influencing job-related characteristics (Priatna & Purnomo, 2020). Consequently, this investigation endeavors to formulate a lecturer performance evaluation instrument tailored to the unique characteristics and requirements of the faculty.

The development of such an evaluation instrument is crucial to ensure a comprehensive assessment of lecturer performance, covering facets indicative of teaching quality, research endeavors, and community service (Retnowati et al., 2017). By employing a robust evaluation instrument, it is anticipated to establish a firm foundation for lecturer career advancement, enhance education quality, and furnish constructive feedback for continuous improvement (Yusrizal, 2017). Similar to prior research by Retnowati et al. (2017) the development of the performance instrument for lecturers comprises four aspects: (1) teaching performance, (2) research performance, (3) community service performance, and (4) teacher capacity.

Presently, the Department of Sport Science at Padang State University lacks a tool for evaluating lecturer performance, a concern prompting researchers to create a performance instrument for assessing teachers at the Faculty of Sport Science, Padang State University. This study will delineate the steps involved in developing a lecturer performance evaluation instrument, encompassing design, validation, and implementation stages (Arthur, 2009). While there have been previous efforts related to the development of lecturer performance instruments (Retnowati et al., 2017; Yusrizal, 2017; Mashud et al., 2018), such endeavors

within the sports context, particularly within the National Faculty of Sport Sciences, are yet to be explored. Consequently, the outcomes of this research are anticipated to positively contribute to enhancing the educational quality at the National Faculty of Sport Science.

METHODS

This study falls within the category of research and development (R&D). Research and Development is a methodical exploration of the design, development, and evaluation process, aimed at establishing an empirical foundation for generating new or enhanced instructional and non-instructional products, tools, and models (Richey et al., 2007). For this research, the employed development design is a procedural model, incorporating various steps essential in the endeavor to create standardized instrument products. The developmental steps or procedures in this study encompass a total of 10 steps, namely:

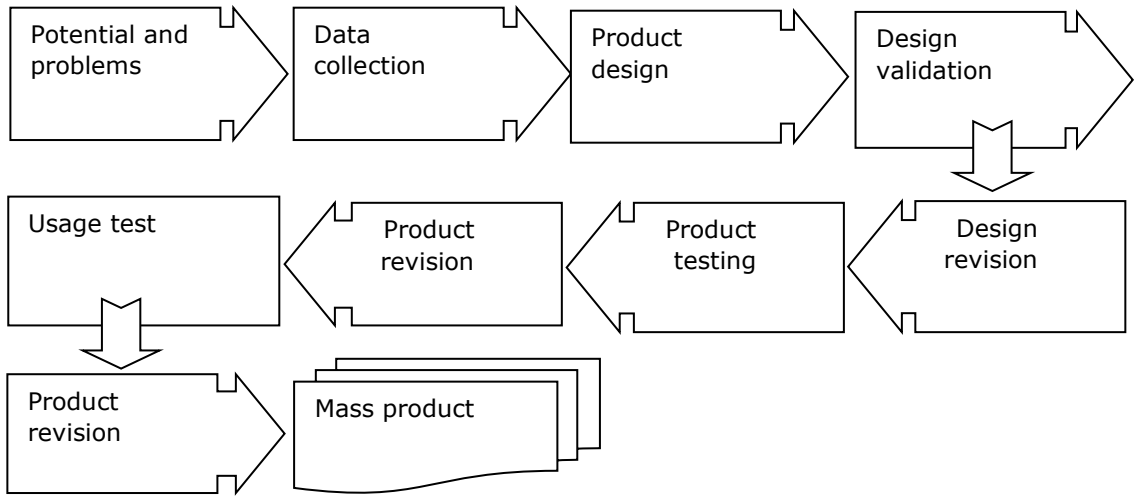

Figure 1: Development chart (Sugiyono, 2016).

The evaluation instrument was developed with a grid consisting of objectives, indicators, sub-indicators and item numbers. The number of each item is, Professional 15 items, social 15 items and personal 15 items. The grids are as follows:

Format 1. Lecturer performance evaluation instrument grid

No	Target	Indicator	Sub Indicator	Item No.	Amount
1	Professional	Quantity of work	The amount of work performed in a specified period of time	1,2,3	3
		Quality of work	Quality of work achieved based on	4,5,6	3

			suitability and readiness requirements		
		Job Knowledge	In-depth knowledge of careers and skills	7,8,9	3
		Creativeness	Originality of ideas arises from actions aimed at solving an emerging problem	10,11,12	3
		Initiative	Passionate about taking on new tasks and increasing responsibility	13,14,15	3
			Amount		15
2	Sosial	Cooperation	Willingness to work with others (organization members)	1,2,3,4,5,6,7,8	8
		Dependability	Perception and reliability of attendance and timely completion of work	9,10,11,12,13,14,15	7
			Amount		15
3	Pribadi	Personal Qualities	It's about personality,	1,3,4	4
			About leadership	5,6,7	3
			Concerning hospitality	8,9,10,11	4
			Concerning personal integrity	12,13,14,15	4
			Amount		15
			Total		45

Product Trial

This research will go through the first product trial stage, namely small group trials. Based on this test data, evaluation and improvement of the product will be carried out. Based on the evaluation and improvement carried out by involving experts, then a large group trial is carried out.

Test Subjects

The test subjects in this study were FIK UNP lecturers in Padang City. Based on data obtained from UNP, the total number of lecturers is 140 people. In the first stage of empirical testing, 30 of the total population will be tested. The second stage will be conducted on 110 people from the total population.

Types of Data

The types of data collected in this study are primary and secondary data. Primary data is data taken directly to the test subjects, namely FIK UNP lecturers. This data is in the form of the results of the trial of the lecturer performance test instrument which was carried out 2 (two) times, a small group of 30 people. For the large group trial, 82 data were collected.

RESULTS AND DISCUSSION

Research Results

The research product that has been designed is then validated so that it can be known whether this product is feasible or not to use. This product validation involved 4 experts; language, design and evaluation, all of whom were UNP experts. The indicators validated consisted of language aspects, for the research product, namely the lecturer performance evaluation instrument, were word choice, ease of understanding the language, and writing style. Aspects validated by design experts are related to aspects of cover design and content design. The clear results can be seen in table 1.

Table 1. Expert assessment of the validity of lecturer performance instruments

No.	Validator	Interrater validity (median)	Description
1	Prof. Dr. Arnedral, M.Pd.	4	Valid
2	Prof. Dr. Syahrastani	4	Valid
3	Prof. Dr. Tjung Haw Sin	4	Valid

Based on the table above, it can be explained that: based on the expert validity test, the median value is obtained in the good category.

The data obtained, expert assessment of the suitability of indicator aspects with sub-indicator aspects, and aspects of research questions, and aspects of the language developed, further carried out interrater reliability tests, with a reliable coefficient value above 0.70 categorized as good. Based on table 2 that, the reliability test of the lecturer performance evaluation instrument Cronbach's alpha is 0.821 can be said to be reliable because it is greater than 0.05, so it can be interpreted that the lecturer performance evaluation instrument has reliability in the high category.

The SPSS out put table above, Intaraclass Correlation Coefficient obtained a Correlation Coefficient value of 0.821. From the data it can be interpreted that the experts have a high level of consistency in providing an assessment of the validity of the lecturer performance evaluation instrument.

Table 2. Reliability of lecturer performance evaluation instruments

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.821	.821	2

Table 3. Intaraclass correlation coefisient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df 1	df 2	Sig
Single Measures	.697[a]	.454	.843	5.591	29	29	.000
Average Measures	.821[c]	.624	.915	5.591	29	29	.000

Limited Trial Data

Based on the results of the analysis of 30 lecturers who were limited test respondents for the lecturer professional performance evaluation instrument, obtained the value of r count = 0.516 > r table = 0.361, because r count is greater than r table for 5% error (0.516 > 0.361), it can be concluded that the professional performance instrument is valid and can be used. For more details can be seen in table 4.

From table 6 that from the reliability test of the professional lecturer performance test instrument, Cronbach's alpha is 0.681, it can be said to be reliable because it is greater than 0.05, so it can be interpreted that the professional lecturer performance instrument has reliability in the sufficient category. Furthermore, the Intaraclass Correlation Coefisient test can be presented in the following SPSS 24 out put.

The SPSS out put table above, Intaraclass Correlation Coefisient presents that the Correlation Coefisient value is 0.672, thus it can be interpreted that the first and second tests have a high level of consistency in carrying out the tests given to the validity of the lecturer's professional performance instrument.

Table 4. Validity of the professional lecturer performance instrument

Korelasi	R Hitung	R Tabel = 0.05	Sig.	Keterangan
X_1 dengan X_2	0.516	0.361	0.003	Valid

*. Correlation is significant at the 0.05 level (2-tailed).

Table 5. Reliability test of professional lecturer performance

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.681	.681	2

Table 6. Intraclass correlation coefficient

Result	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.505[a]	.194	.727	3.136	29	29	.001
Average Measures	.672[c]	.325	.842	3.136	29	29	.001

Based on the results of the analysis of 30 lecturers who were limited test respondents for the lecturer's social performance evaluation instrument, the value of r count = 0.776 > r table = 0.361 was obtained, because r count is greater than r table for 5% error (0.776 >0.361), it can be concluded that the social performance instrument is valid and can be used. For more details can be seen in table 7.

From table 8 that from the reliability test of the lecturer's social test instrument Cronbach's alpha is 0.865 can be said to be reliable because it is greater than 0.05, so it can be interpreted that the lecturer's social performance instrument has reliabelits in the good category. Furthermore, the Intaraclass Correlation Coefisient test can be presented in the following SPSS 24 out put;

The SPSS out put table above, Intaraclass Correlation Coefisient presents that the Correlation Coefisient value is 0.672, thus it can be interpreted that the first and second tests have a high level of consistency in carrying out the tests given to the validity of the lecturer's social performance instrument.

Table 7. Validity of Social Instrument of lecturer performance

Korelasi	R		Sig.	Keterangan
	Hitung	Tabel = 0,05		
X_1 dengan X_2	0,776	0,361	0,000	Valid

Table 8. Social Reliability Test of lecturer performance (Reliability Statistics)

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.865	.865	2

Table 9. Intraclass Correlation Coefficient

Result	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df 2	Sig
Single Measures	.767[a]	.565	.882	7.381	29	29	.000

Average Measures	.868[c]	.722	.937	7.381	29	29	.000

Based on the results of the analysis of 30 lecturers who were limited test respondents for the lecturer's personal performance evaluation instrument, the value of r count = 0.684 > r table = 0.361 was obtained, because r count is greater than r table for 5% error (0.684 > 0.361), it can be concluded that the personal instrument of lecturer performance is valid and can be used. For more details can be seen in table 10.

From table 11 that from the reliability test of the lecturer's social test instrument, Cronbach's alpha is 0.812, it can be said to be reliable because it is greater than 0.05, so it can be interpreted that the lecturer's personal performance instrument has reliability in the good category.

The SPSS out put table above, Intaraclass Correlation Coefisient presents that the Correlation Coefisient value is 0.672, thus it can be interpreted that the first and second tests have a high level of consistency in carrying out the tests given to the validity of the lecturer's personal performance instrument.

Table 10. Validity of Personal Instrument of lecturer performance

Korelasi	**R**		**Sig.**	**Keterangan**
	Hitung	**Tabel = 0,05**		
X_1 dengan X_2	0.684	0.361	0.000	Valid

Table 11. Personal reliability test of lecturer performance (reliability statistics)

Cronbach's Alpha	**Cronbach's Alpha Based on Standardized Items**	**N of Items**
.812	.803	2

Table 12. Intraclass correlation coefficient

Result	**Intraclass Correlation[b]**	**95% Confidence Interval**		**F Test with True Value 0**			
		Lower Bound	**Upper Bound**	**Value**	**df 1**	**df 2**	**Sig**
Single Measures	.670[a]	.417	.827	5.327	29	29	.000
Average Measures	.802[c]	.588	.906	5.327	29	29	.000

Extensive Trial Data

Based on the analysis of 30 lecturers who were limited test respondents for the lecturer professional performance evaluation instrument, the value of r count = 0.911 > r table = 0.214 was obtained, because r count is greater than r table

for 5% error (0.516 > 0.214), it can be concluded that the professional performance instrument is valid and can be used.

From table 14 that from the reliability test of the professional performance test instrument lecturer Cronbach's alpha is 0.901 can be said to be reliable because it is greater than 0.05, so it can be interpreted that the professional performance instrument lecturer has reliabelits in the sufficient category.

The SPSS out put table above, Intaraclass Correlation Coefisient presents that the Correlation Coefisient value is 0.869, thus it can be interpreted that the first and second tests have a high level of consistency in carrying out the tests given to the validity of the lecturer's professional performance instrument.

Table 13. Validity of professional instrument of lecturer performance

Korelasi	R		Sig.	Keterangan
	Hitung	Tabel = 0,05		
X_1 dengan X_2	0.911	0.214	0.000	Valid

Table 14. Reliability test of professional lecturer performance (reliability statistics)

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.901	.901	2

Table 15. Intraclass correlation coefficient

Result	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.768[a]	.520	.876	10.075	81	81	.000
Average Measures	.869[c]	.684	.934	10.075	81	81	.000

Based on the results of the analysis of 82 lecturers who were limited test respondents for the lecturer's social performance evaluation instrument, the value of r count = 0.929 > r table = 0.214 was obtained, because r count is greater than r table for 5% error (0.929 > 0.214), it can be concluded that the social performance instrument is valid and can be used. For more details can be seen in table 16.

From table 17 that from the reliability test of the lecturer's social test instrument, Cronbach's alpha is 0.958, it can be said to be reliable because it is

greater than 0.05, so it can be interpreted that the lecturer's social performance instrument has reliability in the good category.

The SPSS out put table above, Intaraclass Correlation Coefisient presents that the Correlation Coefisient value is 0.958, thus it can be interpreted that the first and second tests have a high level of consistency in carrying out the tests given to the validity of the lecturer's social performance instrument.

Table 16: Validity of the social instrument of lecturer performance

Korelasi	R		Sig.	Keterangan
	Hitung	Tabel = 0,05		
X_1 dengan X_2	0,929	0,214	0,000	Valid

Table 17. Social reliability test of lecturer performance (Reliability Statistics)

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.958	.958	2

Table 18. Intraclass correlation coefficient

Result	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df 2	Sig
Single Measures	.919[a]	.877	.947	23.668	81	81	.000
Average Measures	.958[c]	.934	.973	23.668	81	81	.000

Based on the results of the analysis of 82 lecturers who were limited test respondents for the lecturer's personal performance evaluation instrument, the value of r count = 0.684 > r table = 0.214 was obtained, because r count is greater than r table for a 5% error (0.911 > 0.214), it can be concluded that the personal performance instrument is valid and can be used. For more details can be seen in table 19.

From table 20 that from the reliability test of the lecturer's social test instrument, Cronbach's alpha is 0.953, it can be said to be reliable because it is greater than 0.05, so it can be interpreted that the lecturer's personal performance instrument has reliability in the good category.

The SPSS out put table above, Intaraclass Correlation Coefisient presents that the Correlation Coefisient value is 0.672, thus it can be interpreted that the first and second tests have a high level of consistency in carrying out the tests given to the validity of the lecturer's personal performance instrument.

Table 19. Validity of Personal Instrument of lecturer performance

Korelasi	R		Sig.	Keterangan
	Hitung	Tabel = 0,05		
X_1 dengan X_2	0.929	0.214	0.000	Valid

Table 20. Personal reliability test of lecturer performance (Reliability Statistics)

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.953	.952	2

Tabel 21. Intraclass correlation coefficient

Result	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.910[a]	.864	.941	21.258	81	81	.000
Average Measures	.953[c]	.927	.970	21.258	81	81	.000

Discussion

The study outcomes consist of three components, namely (1) Professional, comprising a total of 15 items, (2) Social, also totaling 15 items, and (3) Personal, which also adds up to 15 items. These three components of the FIK lecturer performance evaluation instrument underwent practicality testing, achieving an average score of 88%, indicating a high level of practicality. Additionally, the effectiveness test yielded an average score of 84.98, rounded up to 85, signifying the high effectiveness of this lecturer performance evaluation instrument. Consequently, throughout the development of this instrument, all valid methods have been employed to create an evaluation tool for assessing the competence of lecturer work (Missaghi-Wedefalk et al., 2012).

Previous research by (Retnowati et al., 2017), followed similar steps in constructing a lecturer performance assessment instrument focused on professional competence. Such competency assessment instruments for teachers demonstrate the correlation between lecturer performance and student learning outcomes (Vogt & Rogalla, 2009). These evaluation instruments serve as a foundation for describing the quality of lecturers and are employed to measure mastery and professional expertise across various subjects and educational levels in Indonesia (Bakri & Budi, 2015; Sumaryanta et al., 2018).

To accurately gauge various competencies of lecturers, competency test instruments must possess ecological and consequential validity (Portelli et al., 2005). Hence, it is imperative to have a lecturer performance evaluation instrument that encompasses professional, social, and personal aspects. Analysis indicates that a lecturer's personal and social attributes have a tangible impact on student learning achievement (Corcoran & Tormey, 2010).

CONCLUSION

The conclusions in this study are; First: the process of developing lecturer performance evaluation instruments at the Faculty of Sport Science starting from the aspects of indicators and sub-indicators, aspects of the presentation of each statement, and aspects of the time used by the instrument itself. These three aspects have been explained well in the instrument that has been developed, and can be used comprehensively in measuring lecturer performance. Second: This research has produced a valid and reliable lecturer performance evaluation instrument at the Faculty of Sport Science. Thirdly, the creation of a pragmatic lecturer performance evaluation instrument for the Faculty of Sport Science. Fourthly, the development of an efficient tool for assessing the performance of lecturers within the Faculty of Sports Science. Future research suggestions may involve the application of the designed instrument to evaluate lecturer performance.

ACKNOWLEDGMENTS

Thanks to the co-authors so that this manuscript is well completed. A big thank you to Padang State University.

REFERENCE

Arthur, L. (2009). From Performativity to Professionalism: Lecturers' Responses to Student Feedback. *Teaching in Higher Education*, *14*(4). https://doi.org/10.1080/13562510903050228

Aziz, I., Okilanda, A., Permadi, A. A., Tjahyanto, T., Prabowo, T. A., Rozi, M. F., Suganda, M. A., & Suryadi, D. (2023a). Correlational Study: Sports Students' Special Test Results and Basic Athletic Training Learning Outcomes. *Retos*, *49*, 519–524. https://doi.org/10.47197/retos.v49.98820

Aziz, I., Okilanda, A., Rozi, M. F., Suganda, M. A., & Suryadi, D. (2023b). Results of Special Tests on Sports Students: Does It Have a Relationship with Learning Outcomes of Basic Athletic Practice? *International Journal of Human Movement and Sports Sciences*, *11*(3), 676–682. https://doi.org/0.13189/saj.2023.110322

Bakri, F., & Budi, S. (2015). Analisis Hasil Uji Kompetensi Guru Fisika. *Jurnal Penelitian & Pengembangan Pendidikan Fisika. 1*(1), 91-96 https://doi.org/10.21009/1.01113

Bangun, S. Y., Sunarno, A., Damanik, S. A., Ilham, Z., & Suganda, M. A. (2023). Modification of Rhythmic Activity Teaching Materials: Study of Development Based on KKNI Curriculum. *Journal of Higher Education Theory and Practice*, *23*(6), 205–215. https://doi.org/10.33423/jhetp.v23i6.5968

Corcoran, R., & Tormey, R. (2010). Teacher Education, Emotional Competencies and Development Education. *Procedia - Social and Behavioral Sciences. 2*(2), 2448-2457. https://doi.org/10.1016/j.sbspro.2010.03.352

Mahmud, H. (2018). Pelaksanaan Model Pengembangan Kinerja Guru pada SMA Negeri di Kota Palopo. *Kelola: Journal of Islamic Education Management*, *1*(1), 1-32. https://doi.org/10.24256/kelola.v1i1.426

Mashud, M., Tangkudung, J., & Widiastuti, W. (2018). Swimming Lesson Based on Interactive Multimedia. *International Journal of Sports Science*, *8*(3), 91–96. https://doi.org/10.5923/j.sports.20180803.04

Mashud, M., Warni, H., Arifin, S., Ferry, M., Pebriyandi, P., & Kristiyandaru, A. (2021). The Application of Discord as an Effort to Increase Students' Wellbeing in Physical Education Learning during the COVID-19 Emergency. *Journal Sport Area*, *6*(3), 335–348. https://doi.org/10.25299/sportarea.2021.vol6(3).6612

Missaghi-Wedefalk, M., Lindh, M., Schön-Ohlsson, C., & Willén, C. (2012). Further Methodological Development of the Test Instrument for Profile of Physical Ability (TIPPA) Designed for Patients with Long-Term Musculoskeletal Pain. *Advances in Physiotherapy*, *14*(3). 97-106 https://doi.org/10.3109/14038196.2012.694907

Nygaard, C., & Belluigi, D. Z. (2011). A Proposed Methodology for Contextualised Evaluation in Higher Education. *Assessment and Evaluation in Higher Education*, *36*(6), 657-671. https://doi.org/10.1080/02602931003650037

Portelli, J. P., Solomon, R. P., Barrett, S., & Mujawamariya, D. (2005). Standardized Teacher Testing Fails Excellence and Validity Tests. *Teaching Education,* *16*(4), 281-295. https://doi.org/10.1080/10476210500345482

Priatna, W., & Purnomo, R. (2020). Implementasi Fuzzy Inference System Metode Sugeno pada Aplikasi Penilaian Kinerja Dosen. *Jurnal Techno.Com*, *19*(3), 245–261. https://doi.org/10.33633/tc.v19i3.3638

Ramli, R., & Jalinus, N. (2013). Evaluasi Kinerja Guru Sekolah Menengah Kejuruan Sumatera Barat Pascasertifikasi. *Jurnal Penelitian dan Evaluasi Pendidikan*, *17*(1), 72-87. https://doi.org/10.21831/pep.v17i1.1362

Retnowati, T. H., Mardapi, D., Kartowagiran, B., & Suranto, S. (2017). Model Evaluasi Kinerja Dosen: Pengembangan Instrumen untuk Mengevaluasi Kinerja Dosen. *Jurnal Penelitian dan Evaluasi Pendidikan*, *21*(2), 206-214. https://doi.org/10.21831/pep.v21i2.16626

Richey, Rita C. & Klein, J. D. (2007). *Design and Development Research: Method, Strategies, and Issues, Manwah, New Jersey:Lawrence Erlbaum Associates, Publisher.*

Sugiyono. (2016). *Metode Penelitian dan Pengembangan (Research and Development)*. Alfabeta.

Sumaryanta, Mardapi, D., Sugiman, & Herawan, T. (2018). Assessing Teacher Competence and Its Follow-up to Support Professional Development Sustainability. *Journal of Teacher Education for Sustainability, 20*(1), 106-123. https://doi.org/10.2478/jtes-2018-0007

Vogt, F., & Rogalla, M. (2009). Developing Adaptive Teaching Competency Through Coaching. *Teaching and Teacher Education, 25*(8), 1051-1060. https://doi.org/10.1016/j.tate.2009.04.002

Widaningrum, I. (2016). Sistem Informasi Evaluasi Kinerja Dosen Aspek Kompetensi dan Tridarma di Lingkungan Universitas Muhammadiyah Ponorogo. *Multitek Indonesia*, *6*(1), 49–60. https://doi.org/10.24269/mtkind.v6i1.121

Wulandari, A., Makhtuna, W., & Mashud. (2023). Dasar Kemampuan Mengajar: Analisis Deskriptif pada Mahasiswa PGSD sebagai Bahan Tindak Lanjut Perkuliahan. *Jurnal Pembelajaran dan Pengajaran Pendidikan Dasar*, *6*(1), 50–59. https://doi.org/10.33369/dikdas.v6i1.26860

Yusrizal, Y. (2017). Pengembangan Instrumen Penilaian Kinerja Dosen di Perguruan Tinggi. *PARAMETER: Jurnal Pendidikan Universitas Negeri Jakarta*, *29*(1), 109-121. https://doi.org/10.21009/parameter.291.10