Desarrollo de instrumentos de evaluación del rendimiento arbitral en deportes de atletismo

Development of referee performance assessment instruments for athletic sports

Authors

Zikrur Rahmat [1]
Heny Setyawati [1]
Donny Wira Yudha Kusuma [1]
Rumini [1]

[1] Universitas Negeri Semarang (Indonesia)

Corresponding author:
Heny Setyawati
henysetyawati@mail.unnes.ac.id

How to cite in APA

Zikrur Rahmat, Heny Setyawati, Donny Wira Yudha Kusuma, & Rumini. (2025). Development of referee performance assessment instruments for athletic sports. *Retos*, *63*, 633–643. https://doi.org/10.47197/retos.v63.111119

Abstract

Introduction: The background of this study is that there is no standardized instrument to measure referees' abilities, so it is urgent to do research. In the future, with the existence of this instrument, it can be an evaluation material and benchmark for referee performance.
Objective: The purpose of development research is to create an athletic referee performance assessment instrument that includes Aspects of Professional Responsibility, Aspects of Rule Mastery, Aspects of Leading Skills, and Aspects of Neatness.
Methodology: The research method used is development research, following 10 research steps with the Research and Development method. The research subjects were athletic referees, as many as 33 people who served in all race numbers. The data analysis technique used is quantitative analysis, which is an assessment using numbers by conducting validity tests, reliability tests, and average analyses.
Results: The developed product's results have advantages, such as the absence of a concrete instrument to measure athletic referees' performance. In addition, it makes it easy to record athletic referees' performance in all athletic match numbers, and the instrument is available in a web-based application version.
Discussion: This instrument is suitable for measuring the performance of athletic referees in running, walking, jumping, and throwing numbers and has been supported by previous research results.
Conclusion: The results of this study conclude that the athletic referee performance assessment instrument that has been developed is feasible to use after being declared a valid and reliable instrument.

Keywords

Athletic, referee, assessment, instrument.

Resumen

Introducción: El antecedente de este estudio es que no existe un instrumento estandarizado para medir la habilidad de los árbitros por lo que es urgente hacer una investigación y en un futuro con la existencia de este instrumento pueda ser un material de evaluación y punto de referencia para el desempeño arbitral.
Objetivo: El objetivo del desarrollo de la investigación es crear un instrumento de evaluación del desempeño de los árbitros de atletismo que incluya Aspectos de Responsabilidad Profesional, Aspectos de Dominio de Reglas, Aspectos de Habilidad para Dirigir, Aspectos de Pulcritud.
Metodología: El método de investigación utilizado es la investigación de desarrollo, siguiendo 10 pasos de investigación con el método de Investigación y Desarrollo. Los sujetos de la investigación fueron árbitros de atletismo, hasta 33 personas que actuaron en todos los números de carrera. La técnica de análisis de datos utilizada es la técnica de análisis cuantitativo, que es una evaluación utilizando números mediante la realización de pruebas de validez, pruebas de fiabilidad y análisis de promedios.
Resultados: Los resultados del producto desarrollado presentan ventajas como la ausencia de un instrumento concreto para medir el rendimiento de los árbitros de atletismo. Además, también facilita el registro del rendimiento de los árbitros de atletismo en todos los números de los partidos de atletismo y el instrumento está disponible en versión de aplicación web.
Discusión: Este instrumento es adecuado para medir el rendimiento de los árbitros de atletismo en los números de carrera, marcha, salto y lanzamiento y ha sido respaldado por los resultados de investigaciones anteriores.
Conclusiones: La conclusión de los resultados de este estudio es que el instrumento de evaluación del rendimiento de los árbitros de atletismo que se ha desarrollado es factible de utilizar tras ser declarado instrumento válido y fiable

Palabras clave

Atletismo, árbitro, evaluación, instrumento.

Introduction

A referee leading a competition is no longer unusual because a referee already understands the rules and has sufficient requirements to become a referee. If a referee does not meet the requirements and does not understand the rules then the referee cannot lead the competition and is not a referee, without a referee, a competition cannot be conducted. The referee's performance in officiating a competition is very great his role is like a king in the competition, whatever the referee's decision in the competition cannot be contested by any party (Loureiro et al., 2023). From the results of the study, researchers found several problems that need to be researched, namely the paradigm of performance assessment instruments (Performance Assessment) is more focused on providing instrument development that has not been found, the absence of standardized instruments to measure referee skills is urgent to do research. Whereas Referee Performance Assessment skills can form intellectual and further develop referee knowledge or concepts. For now, there has been no research specifically with the Athletic Referee Performance Assessment Instrument, but in other sports such as soccer and volleyball there are already, but of course it has quite a difference for athletic referees, the advantage of this Referee Performance Assessment Instrument is that this research will later produce a performance assessment instrument (Performance Assessment) from the above study, the researcher hopes to develop an athletic referee assessment instrument which is considered important and necessary to make an instrument that has a level of validity, reliability and objectivity so that it can be used to assess athletic referees when leading the race.

The referee in officiating the competition must not be arbitrary in giving decisions must be according to the refereeing code of ethics, must not take one side, the referee in officiating the competition must be neutral, fair, and stand on the line of truth seek the truth, a referee is likened to a king in a competition you still have to respect the athlete(Johansen et al., 2018), if an athlete protests against the referee's decision and the referee must respect the athlete and give a clear and firm explanation, he must not be long-winded so that the race continues until the end, likewise, if there is chaotic cheering from the audience, the referee must Respond calmly so that the competition runs successfully. The referee must be ethical, gentle, and firm in leading the race so that the race runs perfectly(Daugherty et al., 2020), with the emergence of this research which aims to develop a standard instrument to measure the performance of athletic sports referees and can be used nationally in general. The existence of a standard instrument for the performance of athletic sports referees is quite important to be able to continue to improve the quality of the implementation of good, efficient, and transparent athletic competitions.

Background

Numerous issues arise for referees in the field during an athletic competition, such as running numbers. Referees sometimes make mistakes when making decisions based on observation, and some athletes turn the wrong way as they enter the field. Referees also oversee the running results of runners on tracks 1, 2, and so forth, often make mistakes when seeing the arrival of runners on lanes 1, 2, and so on, the running referee is also instructed to have a sharp eye, including when starting, making mistakes when giving cards, making mistakes when counting the circumference, forgetting to record the number of times the runner has run around the field, then The jury referee is hesitant to raise the flag when an athlete passes another athlete, in relay events when the baton is passed regardless of the referee's supervision, in jumping events, the same is not free from many mistakes made by referees, including the referee often doubts when an athlete jumps on the diving board whether it passes(Nevill et al., 2013).

Whether or not there is a line, doubts often arise after the results are announced. Observations are also made when athletes jump in the jumping sand pit and the referee's settings are not right. Inexperienced referees also make mistakes because the throw numbers are almost the same, but the referees still make many mistakes, one of which is that the referee often hesitates when the athlete throws even though his feet have passed the throw area(Peng et al., 2021). When the results are revealed, the athlete's observations during the throw, the officials' settings are inappropriate, mistakes are made by inexperienced referees, and the fall of the hammer shot javelin is frequently assessed inaccurately or erroneously, among other reasons, concerns frequently arise. Additionally, since referees frequently make mistakes when making observations, the road number is the same. The

referee must learn all the numbers one day prior to the competition. Other examples of problems that never seem to have an answer include: athletes who walk with their legs bent and float unsteadily when handing out cards; officers who are frequently positioned incorrectly at each post; penalty zones where suggestions from judges or referees have entered the penalty zone(Birinci et al., 2014). The new referee is given the assignment in which number and must be ready even though he has never served in that post and athletics competitions in 1 year can be up to 2 times, especially for Sumatra, different on the island of Java. Athletic competitions are often held.

The challenge of the race in the field number is how the referee's vision when the athlete makes a prefix in the jump and throw, sometimes the prefix in moving there is an error in the vision of the athlete's feet in the throwing sector or the jumping board. Then how to see the fall of the tool in the disc throwing number which sometimes does not cause a sign or mark of the fall of the disc. This is a challenge in the placement of officers who assist the referee in the discus throwing number. Furthermore, the challenge in the javelin throwing number is that sometimes the fall of the javelin also does not cause marks or marks on the field so this becomes a challenge for the referees in charge of the field number and causes confusion for some officers, for example when measuring the fall of the javelin. For the challenge of race referees in track numbers that often occur when athletes are in the corner for running numbers where athletes often enter the trajectory of other runners so that other runners are disturbed, the role of track supervisors is needed to help the track referee become the eyes and ears of the track referee. Every incident of the track supervisor must record it and submit it to the track referee so that the track referee can make decisions for athletes who violate the rules of the race. Furthermore, in other track numbers such as when the athlete makes a false start when giving the baton to the relay run at the time of the limit before and after receiving the relay baton. Based on the explanation of the challenges, more detailed observations and knowledge of the competition rules are needed(Rahmat, 2015)

Current State of the Topic

The problems that exist in an athletic competition are very many for the referee in the field, including the running number the referee makes a decision-making error from the observation that there is a reverse arrival when the athlete enters mixed up, on the timer referee the results of the runner running track 1, 2 and so on, the running number referee also often makes mistakes when looking at the arrival of the runner track 1, 2 and so on, the running referee must also have observant vision at the start, mistaken when giving the card, mistaken when counting around forgetting to record the number of times the runner has run around the field, then the referee jury hesitates to raise the flag when the athlete enters the track of another athlete, in the relay running number when passing the stick regardless of the referee's supervision, in jumping numbers as well, it does not escape the many mistakes made by the referee, including the referee often doubts when athletes jump on the jump board whether they cross the line or not and doubts often occur when announcing results, observations when athletes jump in the sand jumping tub and the arrangement of officers is not appropriate errors occur in inexperienced referees, in throwing numbers it is also almost the same referee still makes many mistakes, namely the referee often doubts when athletes throw even though the athlete's feet have passed the throwing place. doubts often occur when announcing results, doubts often occur at the time of the announcement of the results, observations when athletes when throwing, the arrangement of officers is not appropriate errors occur in inexperienced referees and the fall of the javelin hammer often the measurements are not appropriate or erroneous, and on the road number as well as the referee is often mistaken when looking at observations when athletes walk when their legs are bent and drifting hesitate when giving cards, the placement of officers at each post is often wrong.

From the results of the preliminary study above, researchers found several problems that need to be researched, namely the paradigm of performance assessment instruments (performance assessment) is more focused on providing the development of instruments that have not yet been discovered, the absence of a standard instrument for measuring referees' abilities makes it urgent to carry out research. In fact, referee performance assessment skills can form intellectuals and further develop referee knowledge or concepts(Brymer et al., 2021). license is a requirement to become a referee, while the performance assessment is one of the measuring tools to see the professionalism of a referee in leading a race that must recall the articles and world athletics, the philosophy of the referee, and an

understanding of the rules in accordance with world athletics, an understanding of the division of officers in the race to help the referee so that the race runs smoothly and safely.

The hypothesis that can be examined is that currently there is no specific research on the Athletic Referee Performance Assessment Instrument, but in other sports such as soccer and volleyball there are certainly significant differences for athletic referees(Martinho et al., 2023). The advantage of this Referee Performance Assessment Instrument is that this research will later produce an instrument. performance assessment from the study above, the researcher hopes to be able to develop an instrument for assessing athletic referees which is considered important and needs to create an instrument that has a level of validity, reliability, and objectivity so that it can be used to assess athletic referees when officiating competitions(O'Riordan et al., 2018).

The main purpose of this research is how the License is a requirement to become a referee, while the performance assessment is one of the measuring instruments to see the professionalism of a referee in leading a match that must remember the articles and the world of athletics, the referee's philosophy, and an understanding of the rules by the world of athletics, an understanding of the division of officials in the match to help the referee so that the match runs smoothly and safely with the specification of the Athletic Referee Performance Assessment Instrument. This is one of the products that will be packaged into a book and website

Method

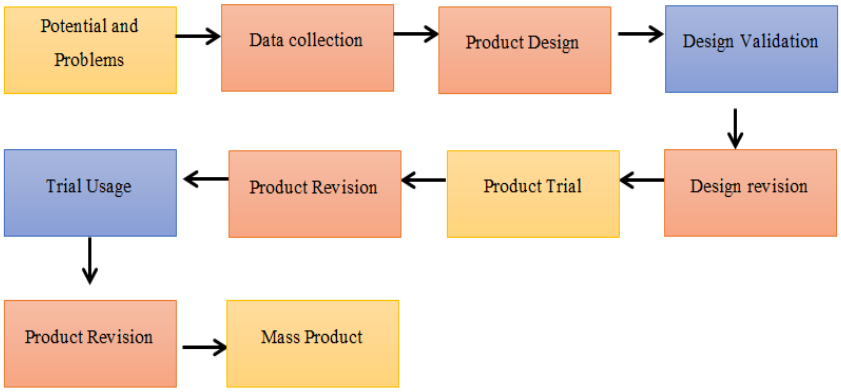
Participants

The subjects in this study were 33 Aceh City Athletics referees, the referees had been and often served at competitions and had attended upgrading and had main, intermediate, and regional licenses.

Procedure

The research model used is the development research model (research and Development or R&D). This model was developed by modifying Borg & Gall's R&D steps.

Figure 1. Steps of Development Research



This research is called research and development and includes the following steps: (1) information gathering and research; (2) planning; (3) product development in initial form; and (4) initial field trial. (5) Main product revision (6) Main field trial (7) Final product revision; (8) Product field trial; (2009) Product operational revision; and (10) Socialization and implementation. Three sports professionals who demonstrated professional responsibility, rule mastery, leadership ability, and neatness completed the validation stage. The participants used the instrument to obtain a stage known as acceptance testing. Retesting to assess the instrument's dependability determined the benefits testing stage

Instrument

States that researchers must follow the steps and instrument development, namely:

1) Defining variables

2) Translating variables into more detailed indicators
3) Compile the grains
4) Conduct a trial
5) Analyzing validity and reliability.

Data analysis

The questionnaire in this research was given to field experts as expert judgment, namely from refereeing 2 people who have intermediate licenses, from coaching who have intermediate licenses, and from the faculty of sports science, to observe the referee who is officiating the competition. The questionnaire answer assessment scale used is(Sugiyono, 2013) stated "There are only two intervals on the Guttman scale, namely Yes and No. Research using the Guttman scale is carried out if you want to get a firm answer to a problem being asked, giving a score of Yes = 1 and no = 0". Using the Guttman scale to ask firm questions and to get firm answers means asking and answering questions should not be long-winded.

Table 1. Research Subjects

No	Research and Development Steps	Technique	Respondent
1	Identification and Potential Problems	Observation	Aceh City Athletics Referee
		Interview	Aceh City Athletics Referee
	Product Validation	Feasibility Test Response Questionnaire by Experts	Media Expert, Test and Measurement Expert, Athletic Expert
	Small Scale Test	Referee Response Questionnaire	30 Referee
	Large Scale Test	Referee Response Questionnaire	50 Referee
		Referee Response Questionnaire	40 Referee

The data analysis technique used in this research is a quantitative analysis technique which is an assessment using numbers. Percentages are intended to determine the status of something that is percentageized and presented in the form of percentages. After obtaining the percentage with the formula, then the feasibility of the Athletic Referee Performance Assessment Instrument (Performance Assessment) in this development research is classified into four categories of feasibility using a Likert scale.

The validity test in instrument development is needed to see whether the measurement instrument used is appropriate for use. The validity test in this study used mean analysis, standard deviation, and factor analysis using SPSS assistance. As it is expected that the assessment instrument must be valid and reliable, then testing the reliability value of the athletic referee performance assessment instrument developed can be done with the test-retest test. So in this case the same instrument is tested, the respondents are the same and the time of implementation is different.

Results

The development of athletic referee performance assessment test instrument developed and has been compiled consists of instruments to measure the performance of athletic referees running numbers, jumping numbers, throwing numbers, and walking numbers. This athletic referee performance assessment instrument was developed to see the competence of the referee from the following angles: 1) Professional responsibility, 2) Rule mastery, 3) Leadership skills, and 4) Neatness.

The product of the development of athletic referee performance instruments is closely related to improving the quality of the referee's performance in carrying out tasks and in developing the referee's self-competence to be better. Improving the referee's performance will also have a positive impact that leads to an increase in the referee's career which will certainly have an impact on the empowerment carried out on referees and other sports human resources. With the presence of this developed athletic referee performance instrument, it is hoped that athletic referees can know and understand the extent of their self-competence as referee professionals in carrying out their duties so far. So that the referees will be able to systematically and significantly improve their competence as an athletic referee profession

The product developed, namely the Athletic Referee Performance Assessment Instrument (Performance Assessment) is an instrument developed to assess the performance of athletic referees/judges in carrying out the task of guarding or leading the race.

1) Referee of athletics running number

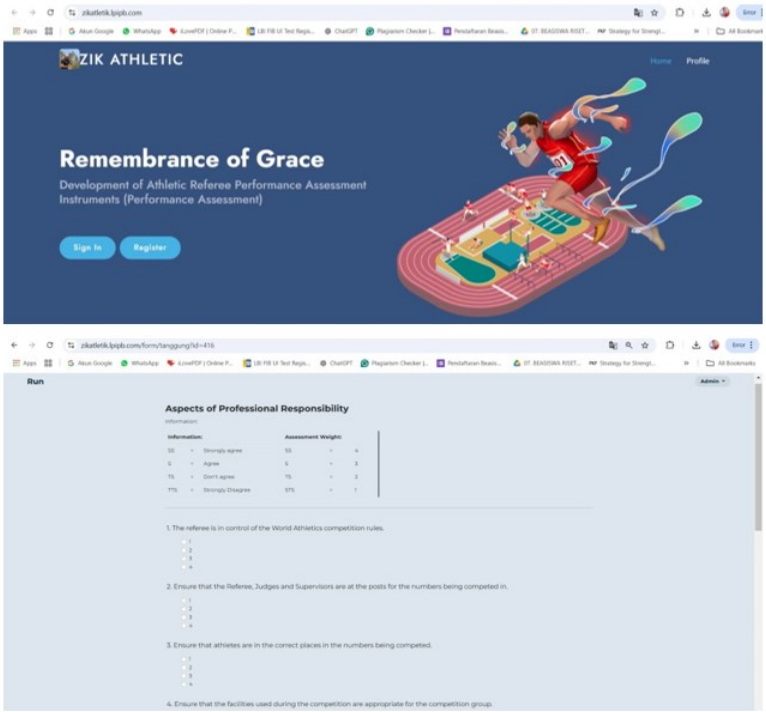
2) Referee of road athletics

3) Referee of athletics jumping numbers
4) Throwing number athletic referee

The developed product instrument is systematically arranged by the stages of development research to analyze the referee's performance against several competencies, namely: 1) the referee's attitude of responsibility towards all components in the race, 2) the referee's mastery of the standard rules in the race, 3) the referee's ability to control, organize, and master the course of the match, and 4) the referee's discipline in terms of neatness of dress. This developed product has advantages such as not yet found concrete instruments to measure the performance of athletic referees. In addition, it also makes it easier to record the performance of athletic referees in all track and field competition numbers and instruments are available in a website-based application version that is easily accessed and used.

Furthermore, in addition to the athletic referee performance assessment instrument product in the form of a manual scoring sheet, it was also developed in an integrated and structured website version and can be accessed on the www.zikatletik.lpipb.com/ page to increase the effectiveness and ease of conducting assessments. The website-based athletic referee performance assessment instrument product can be seen as follows:

Figure 2. Website of athletic referee performance assessment instrument



Analysis of Trial Results

The results of the expert validation test in the form of a questionnaire leaflet regarding the development of athletic referee performance measurement instruments can be seen in table 2.
After receiving expert validator feedback on the Athletic Referee Performance Assessment Instrument, qualitative data was converted to quantitative data using a 5-point rating system, using a scale of 5 shows that the maximum score obtained is 5 and the minimum score is 1. So the ideal mean (Xi) = ½

(5 + 1) = 3, ideal standard deviation (Sbi) = 1/6 (5 - 1) = 0.6. Thus it can be explained that the magnitude of scale 5 = X > 4.08, scale 4 = 3.36 < ≤ 1.92.

Based on the results of the conversion of expert validation value data to the classification score tabulation above, an average score of 4.85 was obtained, which if seen in the classification tabulation is in the very good category. The expert validation results obtained mean that the instrument for measuring the performance of athletic referees that have been prepared has precision and accuracy for measuring the performance of athletic referees which is very good in the eyes of experts.

Table 2. Expert validation test for the development of athletic referee performance measurement instrument

No	Aspects Examined	Indicator	Scoring scale (si)					∑ (fi*si)	Average	Criteria
			1	2	3	4	5			
11	Formulation of indicators for competency achievement	2	0	0	0	0	6	30	5	Very good
22	Grammar	3	0	0	0	0	8	44	4.88	Very good
33	Instrumental Attraction	3	0	0	0	0	5	41	4.55	Very good
44	Ease of Access	2	0	0	0	0	6	30	5	Very good
	Amount	10	0	0	0	5	25	145	19.43	
	Average						4.85			
	Instrument Criteria						Very good			

Table 3 Paired T-Test Test Data for the Running Number Athletic Referee Performance Instrument

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PreTest - PostTest	-.083	6.417	1.852	-4.160	3.994	-.045	11	.965

Referring to the display of the output results above, the sig. value of 0.965> 0.05 is obtained, so according to the basis for decision making in the paired samples T-Test test, it can be concluded that there is no difference between the pretest and post-test data on the development of an athletic referee performance assessment instrument for running numbers.

Table 4 Data of Paired T-Test of Referee Performance Instrument of Athletics Road Number

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PreTest - PostTest	.417	1.832	.529	-.747	1.581	.788	11	.447

Referring to the output results displayed above, the sig. value of 0.447> 0.05 is obtained, so according to the basis for decision making in the paired samples T-test test, it can be concluded that there is no difference between the pretest and posttest data on the development of an athletic referee performance assessment instrument for road numbers.

Table 5 Data of Paired T-Test of Referee Performance Instrument of Athletics Jump Number

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PreTest - PostTest	.750	.754	.218	1.229	.271	3.447	11	.005

Referring to the display of the output results above, the sig. value of 0.05 <0.05 is obtained, so according to the basis for decision making in the paired samples T-test test, it can be concluded that there is no difference between the pretest and posttest test data on the development of an athletic referee performance assessment instrument for jumping numbers.

Referring to the display of the output results above, the sig. value of 0.02 <0.05 is obtained, so according to the basis for decision making in the paired samples T-test test, it can be concluded that

there is a difference between the pretest and posttest data on the development of an athletic referee performance assessment instrument for throwing numbers.

Table 6 Data of Paired T-Test of Referee Performance Instrument of Athletics Throwing Number.

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	PreTest - PostTest	1.333	1.155	.333	2.067	.600	4.000	11	.002

Description of Data Reliability Test Results of Athletic Referee Performance Assessment Instruments

As it is expected that the assessment instrument must be valid and reliable, then testing the reliability value of the athletic referee performance assessment instrument developed can be done with the test-retest test. So in this case the same instrument is tested, the respondents are the same and the time of implementation is different

Table 7 Test-Retest Result Data for Reliability

Pair	Instrument (T-1 dan T-2)	Correlation	Sig.
1	Athletics Running (T1) & Athletics Running (T2)	.991	.000
2	Athletics Road (T1) & Athletics Road (T2)	.957	.000
3	Track & Field (T1) & Jumping (T2)	.994	.000
4	Throwing Number Athletics (T1) & Throwing Number Athletics (T2)	.990	.000

The test-retest results between stages 1 and 2 showed a very good level of reliability for all instruments with a count higher than the r table (0.765). The correlation between the same groups with two treatments showed consistent results, so the instruments used in the first and second-stage trials were declared reliable.

Based on the statistical analysis above, the test-retest output conducted in stages 1 and 2 of the small group trial using all types of instruments showed a very good level of reliability where it was found that the count was greater than the r table. The correlation between the same groups with two treatments showed consistent results so that the measuring instruments developed and used in stages 1 and 2 for small group trials were declared reliable

Discussion

Validity, reliability, objectivity, and norms are the four fundamental elements that must be taken into consideration when creating test instruments. Reliability is the ability of an instrument to deliver the same results when measuring the same object multiple times. Validity is the ability to measure what should be measured(Cunningham & Sullivan, 2020). A crucial phase in the development and assessment of instruments is instrument validation. Validity is the degree to which the test's components are accurate and the degree to which the score makes sense given its interpretation(Purnomo et al., 2024). Collecting data to demonstrate the scientific foundation for interpreting the score by the assessment instrument's stated aim is a part of the validation process. Stated differently, the interpretation of the assessment results' score can be based on the intended use of the tool. But first, you must authenticate the instrument to acquire the correct interpretation(Ngatman et al., 2024).

An instrument can be validated in several methods, most of which fall into one of three categories: constructions, based on criteria, or content. Construct validation involves developing an instrument based on theoretical frameworks and existing literature, ensuring that the instrument accurately reflects the constructs it seeks to measure. Content validation assesses whether the instrument comprehensively covers the domain of interest, often involving expert reviews to evaluate the relevance and clarity of the items. Criterion validation involves comparing the instrument to other established measures to assess its validity. Validation of referee instruments involves a systematic process of ensuring their reliability and validity, which is critical to their credibility and utility in

research or practical applications. This process typically includes both qualitative and quantitative methods to assess various aspects of the instrument, such as content validity, construct validity, and reliability. The given paper offers an insight into the different methodologies and approaches used in the validation of various instruments, which can be applied to the context of referee instruments(Collet et al., 2011). If the assessment instrument's content is thorough, pertinent, and stays within the parameters of the measurement objective, then it is considered legitimate(Morris & O'Connor, 2017). Finding the instrument's content is regarded as legitimate and verifiable based on a logical examination of the information, whose evaluation is dependent on personal, subjective factors(Westbrooks et al., 2023). To account for the results acquired, the people who are deemed experts in the components being tested are the appropriate ones to determine content validity(Weston, 2015)

The referee's ability when leading a race plays a very important role in the success of a race in running, walking, throwing, and jumping sports. The performance of athletic referees must have a standard assessment to determine the referee's ability, with the existence of standard measuring instruments, it will be possible to determine the referee's performance in leading running, walking, throwing, and jumping competitions(Gomes et al., 2022). Based on the level of athletic referee needs, it is necessary to develop a test instrument to assess the performance of athletic referees in running, walking, throwing, and jumping. Thus, it is necessary to study the form of observation in the field when the referee leads the race, in a series of measurements carried out specifically concerning the development of an athletic referee assessment performance instrument(Sinval et al., 2020). Its development refers to the part and whole learning theory. What is meant is that its development includes the following parts; 1) The referee's performance assessment instrument; 2) items/parts of the athletic referee performance assessment instrument developed; 3) The referee's performance assessment instrument from each item/part developed; 4) a combination of websites and books for the athletic referee performance assessment instrument developed(Myers et al., 2012).

There hasn't been any particular research done using the Athletic Referee Performance Assessment Instrument as of yet, but there must be a significant difference for athletic referees in other sports like football and volleyball(Mudian et al., 2021). The soccer referee performance instrument research results state that This study identifies a 7-point definition of successful referee performance and outlines 26 characteristics that referees consider important, as determined by various stakeholders in Major League Soccer and the Professional Referee Organization. Although this paper does not specifically focus on the assessment of referee performance during matches, it provides insight into the qualities that contribute to referee success, which can inform the evaluation of performance in game situations(Carvalho et al., 2023). The benefit of this referee performance assessment tool is that it will eventually yield a performance assessment tool (Performance Assessment) from the aforementioned study. In general, the instrument shows satisfactory results in terms of clarity of language, practical relevance, and theoretical relevance. In this sense, it is worth raising hypotheses as to why certain items, individually, did not obtain the desired results. A first point that allows for reflection is because these items belong to the same dimension - Social Responsibility. It can be assumed that some of the judges' conceptions of the dimension differ from those adopted by the authors. The concept of social responsibility covers both economic-financial, legal, and ethical aspects. In addition, social responsibility is understood as supporting the community in which it operates, promoting the well-being of its employees, fostering synergy between partners, and increasing customer and consumer satisfaction(Furtado et al., 2016).

The novelty of this study is to produce an athletic referee performance instrument, with the resulting product in the form of an instrument that is entered into a website, printed book, and indexed international journal which can later be utilized and help referees well in national athletic competitions. Overall, the results of this study support existing research on athletic referee performance assessment instruments.

Implications The findings and recommendations of this study are that the athletic referee performance assessment instrument has proven feasible and effective for use by referees in athletic competitions. Therefore, it is also prioritized to be able to measure the performance of athletic referees. Then the results of this study also use a website-based platform to help the referee's performance process in the

field and develop an athletic referee performance assessment system that has an important role in creating a positive race climate in various events that take place

This research and development has been attempted and carried out by existing procedures, however, it still has limitations, namely, the research time used in this study is relatively short when in fact, this research still has many ideas that can still be developed and requires longer time and requires large financing.

Conclusions

The athletic referee performance assessment tool is grounded in empirical research. This is further reinforced by the references to earlier research that are included in the discussion of findings and remarks. These results suggest developing a tool to assess athletic referee performance in a competition. The use of referee performance instruments available on a website allows referees to be measured for their level of understanding, knowledge, and ability to participate in a competition. In addition, the results of the study indicate that the effectiveness and efficiency of products in the classroom are ideal for improving athletic referee performance. The results of this study were also refined by expert evaluation. Selection of referees in the future, especially athletics based on performance through assessment instruments. Develop an assessment system framework to evaluate the performance of referees while on duty in the field.

The impact of the referee performance assessment system to make athletic sports referees more professional and in the recruitment of referees who will be assigned to competitions more selectively.

Acknowledgments

We would like to thank the Ministry of Education, Culture, Research and Technology of the Republic of Indonesia through the Higher Education Financing Center (BPPT) and the Education Fund Management Agency (LPDP) for providing financial support/sponsorship for this research.

Financing

We would like to thank the Ministry of Education, Culture, Research and Technology of the Republic of Indonesia through the Higher Education Financing Center (BPPT) and the Education Fund Management Agency (LPDP) for providing financial support/sponsorship for this research.

References

Birinci, M. C., Yılmaz, A. K., Erkin, A., Şahbaz, S., & Aydın, İ. (2014). Determination of Relationship Between Respiratory Parameters and Aerobic Capacity of Referees. *Procedia - Social and Behavioral Sciences*, *152*, 1267–1273. https://doi.org/10.1016/j.sbspro.2014.09.368

Brymer, R., Rodenberg, R. M., Zheng, H., & Holcomb, T. R. (2021). College Football Referee Bias and Sports Betting Impact. *Eastern Economic Journal*, *47*(1), 91–106. https://www.proquest.com/docview/2478672821?sourcetype=Scholarly Journals

Carvalho, V., Esteves, P. T., Nunes, C., Helsen, W. F., & Travassos, B. (2023). The assessment of the match performance of association football referees: Identification of key variables. *PLoS ONE*, *18*(9 September), 1–9. https://doi.org/10.1371/journal.pone.0291917

Collet, C., do Nascimento, J. V., Ramos, V., & Stefanello, J. M. F. (2011). Construção e validação do instrumento de avaliação do desempenho técnico-tático no voleibol. *Revista Brasileira de Cineantropometria e Desempenho Humano*, *13*(1), 43–51. https://doi.org/10.5007/1980-0037.2011v13n1p43

Cunningham, I., & Sullivan, P. (2020). Testing the factor validity of the Referee Self-Efficacy Scale (REFS) in non-invasion sports officials. *International Journal of Sport, Exercise and Health Research*, *4*(1), 12–15. https://doi.org/10.31254/sportmed.4103

Daugherty, J., DePadilla, L., & Sarmiento, K. (2020). Assessment of HEADS UP online training as an educational intervention for sports officials/athletic trainers. *Journal of Safety Research*, *74*, 133–141. https://doi.org/10.1016/j.jsr.2020.04.015

Furtado, S., Treter Gonçalves, G. H. ., & Marinho Mezzadri, F. . (2024). Evidencia de validez de un nuevo instrumento que mide la percepción de (buena) gobernanza en las organizaciones deportivas (Validity evidence for a new instrument measuring the perception of (good) governance in sports organizations). *Retos*, *55*, 624–634. https://doi.org/10.47197/retos.v55.103809

Gomes, A. R., Nogueira, D., Fontes, L., Gomes, A. R., & Resende, R. (2022). Referees' Emotions and Performance Perception: The Importance of Stress and Cognitive Appraisal. *Revistas*, *22*, 156–170. https://doi.org/10.6018/cpd.468751

Johansen, B. T., Ommundsen, Y., & Haugen, T. (2018). Referee efficacy in the context of Norwegian soccer referees – A meaningful construct? *Psychology of Sport and Exercise*, *38*, 184–191. https://doi.org/10.1016/j.psychsport.2018.06.013

Loureiro, M., Mesquita, I., Ramos, A., Coutinho, P., Ribeiro, J., Clemente, F. M., Nakamura, F. Y., & Afonso, J. (2023). Flexible Training Planning Coupled with Flexible Assessment: A 12-Week Randomized Feasibility Study in a Youth Female Volleyball Team. *Children*, *10*(1). https://doi.org/10.3390/children10010029

Martinho, D. V., Field, A., Rebelo, A., Gouveia, É. R., & Sarmento, H. (2023). A Systematic Review of the Physical, Physiological, Nutritional and Anthropometric Profiles of Soccer Referees. *Sports Medicine - Open*, *9*(1). https://doi.org/10.1186/s40798-023-00610-7

Morris, G., & O'Connor, D. (2017). Key attributes of expert NRL referees. *Journal of Sports Sciences*, *35*(9), 852–857. https://doi.org/10.1080/02640414.2016.1194524

Mudian, D., Asmawi, M., Rihatno, T., Lanos, M. E. C., Elisyah, V., & Aji, T. (2021). Trend improving soccer professional referees performance in west java using psychological skill training and physical fitness. *International Journal of Human Movement and Sports Sciences*, *9*(3), 595–601. https://doi.org/10.13189/saj.2021.090328

Myers, N. D., Feltz, D. L., Guillén, F., & Dithurbide, L. (2012). Development of, and Initial Validity Evidence for, the Referee Self-Efficacy Scale: A Multistudy Report. *Journal of Sport and Exercise Psychology*, *34*(6), 737–765. https://doi.org/10.1123/jsep.34.6.737

Nevill, A., Webb, T., & Watts, A. (2013). Improved training of football referees and the decline in home advantage post-WW2. *Psychology of Sport and Exercise*, *14*(2), 220–227. https://doi.org/10.1016/j.psychsport.2012.11.001

Ngatman, N., Hermawan, H. A., Komari, A. ., Solikhin, M. N., & sunaryo, sunaryo. (2024). Modelo de instrumento de evaluación de técnicas básicas de saque, golpe de fondo de derecha y de revés, volea de derecha y de revés basado en el "método de acentos" para tenistas juveniles de la región de Yogyakarta (Model of assessment instrument for basic techniques of serving, forehand and backhand groundstroke, forehand and backhand volley based on "acentos method" for junior tennis players in Yogyakarta region). *Retos*, *59*, 584–594. https://doi.org/10.47197/retos.v59.103223

O'Riordan, C. N., Newell, M., & Flaherty, G. (2018). Cardiovascular disease risk factor profile of male Gaelic Athletic Association sports referees. *Irish Journal of Medical Science (1971 -)*, *187*(4), 915–924. https://doi.org/10.1007/s11845-018-1792-3

Peng, D., Wang, C., Ma, R., Mao, S., Qu, S., Ren, Z., Golovynskyi, S., & Pan, C. (2021). Mechanoluminescent materials for athletic analytics in sports science. *Science Bulletin*, *66*(3), 206–209. https://doi.org/10.1016/j.scib.2020.09.029

Purnomo, E., Ma'mun, A., Winarno, M. E., Mardesia, P., Jermaina, N., & Zainal Abidin, N. E. (2024). Reliability and Interrater Agreement: Development andValidation of a UniversalValues in Sport Instrument. *Retos*, *57*, 664–683. https://doi.org/10.47197/retos.v57.107449

Rahmat, Z. (2015). *Atletik Dasar & Lanjutan,* (Issue 1). Natural Aceh.

Sinval, J., Aragão e Pina, J., Sinval, J., Marôco, J., Santos, C. M., Uitdewilligen, S., Maynard, M. T., & Passos, A. M. (2020). Development of the Referee Shared Mental Models Measure (RSMMM). *Frontiers in Psychology*, *11*(October). https://doi.org/10.3389/fpsyg.2020.550271

Sugiyono, D. (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan Tindakan*. ALFABETA.

Westbrooks, J., Low, D. A., & Brownlee, T. E. (2023). Determining the definition and components of successful soccer referee performance. *Journal of Sports Sciences*, *24*(August), 1–23. https://doi.org/10.1080/02640414.2024.2317641

Weston, M. (2015). Match performances of soccer referees: the role of sports science. *Mov Sport Sci/Sci Mot*, *87*, 113–117. https://doi.org/10.1051/sm/2014011

Authors' and translators' details:

Zikrur Rahmat	zikrur@students.unnes.ac.id	Author
Heny Setyawati	henysetyawati@mail.unnes.ac.id	Author/ correspondiente
Donny Wira Yudha Kusuma	donnywirayudhakusuma@mail.unnes.ac.id	Author
Rumini	rumini@mail.unnes.ac.id	Translator