Jurnal Cakrawala Pendas
Vol. 10 No 4, October 2024
DOI: http://dx.doi.org/10.31949/jcp.v10i411284

p-ISSN: 2442-7470
e-ISSN: 2579-4442

DEVELOPMENT OF TEACHER PERFORMANCE EVALUATION MODELS OF INCLUSIVE AND REGULAR SCHOOLS

Yuyun Yuningsih Ratnasari[1], Nurwalidainismawati [2], Abas Oya[3] , Tri Nurrahmania[4]
[1,2,3]STKIP Harapan Bima
[1]yuyunyuningsihratnasari39@gmail.com

Abstract
The objective of this study is to develop a performance evaluation model for teachers in regular and inclusive schools. The research followed stages of analysis, design, development, implementation, and evaluation. Data were collected through observation, documentation, and questionnaires and analyzed using both descriptive and qualitative methods. The Gregory formula was used to analyze expert validation, and confirmatory factor analysis (CFA) was employed to assess the test results. Findings indicate that the evaluation instrument meets validity criteria, with loading factor values ranging from 0,7 to 1.00 (>0.7) for SDN Doridungga. Meanwhile, SLB Wihdatul Ummah Donggo with loading factor values ranging from -0.074 to 0.979 (>0.7). Confirmatory Factor Analysis (CFA) shows that the model for SDN Doridungga has a good fit with a Standardized Root Mean Square Residual (SRMR) value of 0.001, while the model for SLB Wihdatul Ummah Donggo shows an SRMR value of 0.258, which indicates that the model is not fit. The findings have significant practical implications for teacher evaluation practices. For regular schools like SDN Doridungga, the model proves to be a reliable tool for enhancing teacher effectiveness and improving student outcomes. In contrast, inclusive schools, such as SLB Wihdatul Ummah Donggo, require further adaptation of the instrument to reflect the unique teaching challenges present in these environments.
Keywords: *Evaluation Models; Inclusive and Regular Schools*

Abstrak
Tujuan dari penelitian ini adalah untuk mengembangkan model evaluasi kinerja bagi guru di sekolah reguler dan inklusif. Penelitian mengikuti tahapan analisis, desain, pengembangan, implementasi, dan evaluasi. Data dikumpulkan melalui observasi, dokumentasi, dan kuesioner dan dianalisis menggunakan metode deskriptif dan kualitatif. Rumus Gregory digunakan untuk menganalisis validasi ahli, dan analisis faktor konfirmasi (CFA) digunakan untuk menilai hasil tes. Temuan menunjukkan bahwa instrumen evaluasi memenuhi kriteria validitas, dengan nilai faktor pemuatan berkisar antara 0,7 hingga 1,00 (>0,7) untuk SDN Doridungga. Sementara, SLB Wihdatul Ummah Donggo dengan nilai loading factor berkisar antara -0,074 hingga 0,979 (>0,7). Confirmatory Factor Analysis (CFA) menunjukkan bahwa model untuk SDN Doridungga memiliki fit yang baik dengan nilai Standardized Root Mean Square Residual (SRMR) 0.001, sedangkan model untuk SLB Wihdatul Ummah Donggo menunjukkan nilai SRMR 0.258, yang mengindikasikan bahwa model tersebut tidak fit. Temuan ini memiliki implikasi praktis yang signifikan terhadap praktik evaluasi guru. Untuk sekolah reguler seperti SDN Doridungga, model ini terbukti menjadi alat yang dapat diandalkan untuk meningkatkan efektivitas guru dan meningkatkan hasil belajar siswa. Sebaliknya, sekolah inklusi, seperti SLB Wihdatul Ummah Donggo, membutuhkan adaptasi lebih lanjut dari instrumen tersebut untuk merefleksikan tantangan pengajaran yang unik yang ada di lingkungan ini.
Kata Kunci: Model Evaluasi; Sekolah Inklusif dan Reguler

Received	: 15 September 2024	Approved	: 28 October 2024
Reviews	: 27 Octobe 2024	Published	: 31 October 2024

Introduction

Basic education plays an important role as the basis for shaping students' abilities and character. One of the main factors that affect the quality of basic education is teacher performance. Teacher performance evaluation is a crucial tool in ensuring that the education provided meets the expected standards, as well as one of the important factors for success in the education and training system (Saljooghi & Salehi, 2016), the process of arriving at judgment about individual teacher's performance against the background of his work environment and his future potential for the school system (Chukwubikem, 2013). In Indonesia, the current teacher evaluation system is marked by a mix of traditional and innovative approaches aimed at enhancing teacher quality and, ultimately, student learning outcomes. These evaluations are conducted through mechanisms such as the Teacher Competency Test (UKG) and Teacher Performance Evaluation (PKG), which involve various stakeholders like school principals and teachers. While stakeholders show support for reform and new evaluation schemes, challenges remain, especially in aligning these evaluations with the needs of inclusive education (Perez-Alvarez et al., 2020). However, in the context of basic education, there are challenges in formulating and implementing effective teacher performance evaluation models, especially among inclusive and regular schools.

The main problem faced is the lack of a teacher performance evaluation model that is by the needs and characteristics of basic education in these two types of schools. Inclusive and regular schools have different needs and challenges, so an evaluation framework is needed that can accommodate these differences and ensure a holistic approach that can improve teaching practices in diverse learning environments. Teacher evaluations must respond to the specific problems of the environment in which the teacher works, as well as understand that teaching is a complex activity that needs to be analyzed from multiple perspectives (Gómez López & Valdés, 2019; Guía, 2012).

Inclusive schools face significant challenges in teacher evaluation due to the diverse needs of their student populations, particularly those with special educational needs (SEN), which demand specialized pedagogical strategies. Unfortunately, current evaluation tools like the Marzano Teacher Evaluation System do not adequately address the complexities of inclusive teaching, as they fail to align with the individualized approaches necessary in such settings, leading to dissatisfaction among educators and administrators (Thompson & Lesh, 2023). Moreover, teachers often report feeling overwhelmed due to insufficient training and limited resources, which are essential for managing the diverse demands of inclusive classrooms (Cohitmingao, 2024). The urgency of this research lies in the importance of developing an evaluation model that can meet the specific needs of the two types of schools. Without a proper evaluation model, it isn't easy to ensure that teacher performance can be effectively measured and improved, which can negatively impact the quality of education students receive.

From the review of the existing literature, several key themes emerged related to the development of teacher performance evaluation models for primary education, especially in comparing inclusive and regular schools. First, the role of teachers' unions in influencing education production emphasizes the importance of considering the impact of teachers' unions on public school performance in the framework of evaluation (Mardia & Mukhtar S, 2022). Second, the importance of evaluating teacher performance in an inclusive environment emphasizes the need for a tailored evaluation process to address the specific requirements of an inclusive setting (Hermanto, 2022). This emphasizes the need for a specific evaluation process that can accommodate the various challenges faced in inclusive education. Theoretical frameworks such as dynamic models of educational effectiveness and organizational perspectives are mentioned as important in school effectiveness research, providing valuable insights for designing comprehensive teacher performance evaluation models (Scheerens, 2013). Additionally, the need for evidence-based standards for instructional effectiveness in teacher evaluation is essential to ensure teacher evaluations that are thorough, fair, and aligned with best practices (Permana & Eliza, 2022).

Discussion on the impact of the inclusive education climate of schools on teacher competence, Xue et al. (2023) highlighting the importance of creating a supportive inclusive practice environment to improve teachers' ability to serve the diverse needs of students. Finally, challenges and solutions related to the implementation of inclusive education, including teacher understanding, resource availability, and curriculum support, were raised (Dewi et al., 2020). Addressing these barriers is essential for developing an effective teacher performance evaluation model that considers the unique requirements of inclusive education.

The state of the art of the study includes several related articles. One of them is a study that explores the characteristics of teacher evaluation models in Finland with naturalistic research design and qualitative content analysis. The results show that the evaluation model in Finland prioritizes teacher empowerment and professional development through focused evaluation (Tarhan et al., 2019). Other articles discuss the monitoring and evaluation of teachers' effectiveness using the *Teacher Performance Appraisal and Development tool* (TPAD) in public high schools in Kenya. The research method is a descriptive survey, and the results show that teacher attendance monitoring, teacher class list monitoring, and adherence to school deadlines have a positive effect on teacher performance (Ibrahim, 2020). In addition, there is research that creates a tool to evaluate the effectiveness of teachers in boarding schools. The methods used include planning, instrument trials, and measurements. The results show that the teacher performance assessment instrument in boarding schools meets the validity criteria (Saleh et al., 2024). In contrast, the novelty of this research lies in the development of a teacher performance evaluation model tailored specifically for inclusive and regular schools in the Indonesian context.

This research aims to develop a framework for evaluating teacher performance that is tailored to the needs of inclusive and regular schools. The study integrates an inclusive perspective into the evaluation process, ensuring that the unique challenges of teaching in inclusive settings are addressed. Additionally, the framework emphasizes the adoption of evidence-based evaluation standards and promotes active stakeholder engagement, including teachers, administrators, and education experts. A key focus of the research is on continuous testing and revision of the model to ensure its adaptability and effectiveness. By addressing these aspects, the study seeks to improve the quality of teaching and learning in both inclusive and regular educational environments.

Research Methods

The research method is research and development (R&D) using the ADDIE (Analysis, Design, Development, Implementation, and Evaluation) model research design to develop a teacher performance evaluation model for SDN Doridungga and SLB Wihdatul Ummah Donggo. The development of this teacher performance evaluation model will involve the systematic application of the stages of the ADDIE model to ensure a comprehensive and effective approach. Here is how each stage of the ADDIE model will be elaborated in the context of developing an evaluation model:

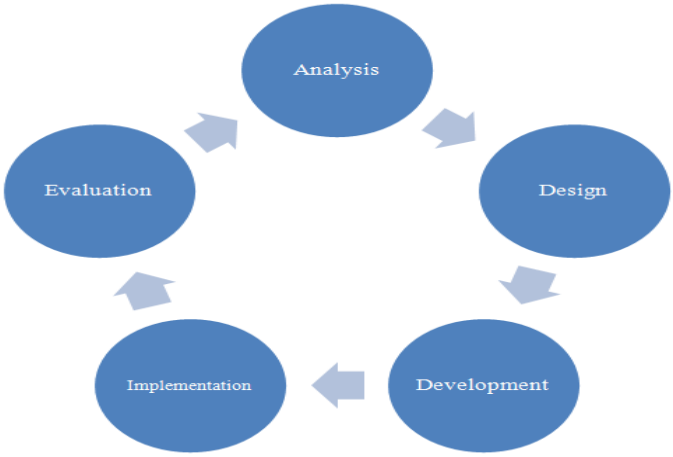

Diagram 1. The ADDIE Model (Oya & Haryanto, 2022)

The first stage carried out is the analysis stage. In this stage, conduct a thorough analysis of current teacher performance evaluation practices at SDN Doridungga and SLB Wihdatul Ummah Donggo, and identify specific requirements and challenges faced by teachers and administrators. The number of PTKs at SDN Doridungga is 22 people. The number of PTK at SLB Wihdatul Ummah Donggo is 12 people. Additionally, understanding the school's unique context, including student demographics, teaching methodologies, and existing evaluation frameworks.

The next stage is the design stage. This stage includes the design of instructional solutions. Based on the findings of the analysis, design a customized teacher performance evaluation model that is in line with the school's educational goals, values, and philosophy. In addition, the development of criteria and metrics defines clear evaluation criteria and performance metrics that reflect the desired outcomes and are in line with the school's educational goals. Research indicates that well-defined metrics, which target instructional practices and student engagement, provide teachers with actionable feedback that promotes continuous professional development (Stronge, 2012).

The next stage is the development stage. At this stage, develop the tools and materials necessary to implement the evaluation model, such as evaluation forms and guidelines for observation. Then, it is followed by conducting a content validity test to assess feasibility and make necessary adjustments before full implementation. The validity test of this content uses a Gregory test.

Then the next stage is the implementation stage. Based on the results of the development, it is followed by conducting implementation, or field tests in regular and inclusive schools to find out the results of the products developed.

The last is the evaluation stage. This stage evaluates the impact of the new evaluation model on teacher performance, student outcomes, and overall school improvement based on predetermined success indicators. Additionally, collect feedback from teachers, students, and administrators to continuously improve and refine the evaluation model for continued effectiveness.

This study uses various data collection techniques, including observation, document analysis, and questionnaires with the Likert scale. Data analysis involves a mixed approach, a research approach that involves the collection of quantitative and qualitative data (Cresswel, 2016). The development of the questionnaire followed a rigorous process to ensure its validity and reliability. Initial drafts of the questionnaire items were based on an extensive literature review and input from educational experts. The items covered key aspects of teacher performance, including classroom management, student engagement, and teaching methods. Gregory's method of content validation was employed to ensure that the questionnaire items were aligned with the evaluation objectives. Expert judgments were sought on the clarity, relevance, and comprehensiveness of each item, and modifications were made accordingly.

A mixed-methods approach was adopted, combining qualitative data from observations, and analysis of documents with quantitative data obtained from Likert scale questionnaires to provide a comprehensive understanding of the developed teacher performance evaluation model. Data analysis was conducted using SmartPLS software to model structural equations and perform quantitative analysis. The use of this application is to analyze the relationship between variables, test the proposed model, and assess the effectiveness of the teacher performance evaluation framework developed using the ADDIE model.

Results and Discussion

Results

Analysis Stage

The analysis process begins with a conceptual and theoretical study to formulate performance indicators, which include three main aspects: classroom management, student engagement, and teaching methods. These indicators were carefully selected to address the unique challenges faced by both regular and inclusive schools. Data is collected through surveys and Focus Group Discussions (FGDs) with teachers, which are then validated to ensure the relevance and accuracy of the indicators set.

A critical consideration during the analysis was the varying contexts of regular and inclusive schools. In regular schools, teachers typically work with a more homogeneous student population, allowing for the implementation of standardized pedagogical strategies. This context tends to facilitate stronger compliance with the evaluation criteria, resulting in higher performance scores. Conversely, teachers in inclusive schools face significant challenges, such as accommodating students with diverse abilities and special educational needs (SEN). These complexities necessitate individualized instructional approaches, which can complicate the application of standardized evaluation indicators and ultimately impact performance scores negatively.

For example, during FGDs, several teachers expressed feeling overwhelmed by the need to adapt their teaching methods for students with varying needs, which sometimes resulted in lower engagement scores on the evaluation. Additionally, certain indicators, like classroom management, may not adequately reflect the unique dynamics of an inclusive classroom where multiple needs must be addressed simultaneously.

To effectively capture these nuances, the analysis emphasized the need for flexible evaluation criteria that acknowledge the specific instructional approaches required in inclusive classrooms. This adaptability is crucial to ensure that the evaluation model can accurately reflect the performance of educators in diverse educational settings. After validation, trials of the instrument are conducted to ensure its effectiveness and suitability before being widely applied in these varied contexts.

Design Stage

This study designs a teacher performance evaluation model that is by the needs of regular and inclusive schools, based on the needs analysis results. This evaluation model covers three main aspects: Classroom Management, Student Engagement, and Teaching Methods. Each aspect is equipped with relevant indicators for both types of schools.

Table 1 shows the aspects and indicators used in the evaluation instrument

NO	Aspects	Indicators
1	Classroom Management	Preparation and preparation of lesson plans by curriculum standards. The teacher's ability to overcome distractions and maintain student concentration in the classroom. The quality of the learning environment supports the safety and comfort of students. Provide appropriate materials and strategies for students with special needs.

Jurnal Cakrawala Pendas Vol. 10, No. 4, October 2024, pp. 914-927

2	Student Engagement	The level of student attendance and their involvement in class activities. The level of enthusiasm and interest of students in the material taught. The seriousness of students in trying to understand and master the material taught. Provide emotional support for students in need.
3	Teaching Methods	The quality of the teacher's explanation in explaining the material in a clear and structured way. The effectiveness of teachers in using questions to stimulate students' critical thinking. The teacher's ability to facilitate class discussions and interactions between students. The ability of teachers to adapt teaching and assessment methods to meet the special needs of students.

Development Stage

This stage of development is the result of revisions based on recommendations from experts (Jugjes). This involves having experts review the items to ensure they are appropriate and comprehensive for the construction being measured (Tenorio & González Ortega, 2024). The assessment by the two Jugjes corresponds to Gregory's theory. The assessment results showed that one indicator needed to be improved, namely indicator 8 on the aspect of student involvement related to language.

Table 2. Showing the results of V Gregory's analysis of the instrument:

Items	**Jugjes 1**		**Jugjes 2**	
	Relevant	**Irrelevant**	**Relevant**	**Irrelevant**
1	√		√	
2	√		√	
3	√		√	
4	√		√	
5	√		√	
6	√		√	
7	√		√	
8	√		√	
9	√		√	
10	√		√	
11	√		√	
12	√		√	

By **Table 2**, all instrument items are declared "relevant" according to V Gregory's criteria. This shows that the instrument items of the teacher performance model have high content validity in terms of their conformity with the indicators (Almanasreh et al., 2019; L'Ecuyer et al., 2020; Shrotryia & Dhanda, 2019; Zamanzadeh et al., 2015). These findings imply that the theoretically developed instrument for assessing teacher performance in regular and inclusive schools is proven to have good content validity. Thus, this teacher performance evaluation tool can be measured accurately and practically in both types of schools.

Implementation Stage

Construct Validity

The data obtained from the limited trial of this teacher performance instrument was analyzed using the SmartPLS application with the second-order Confirmatory Factor Analysis

(CFA) approach. CFA is used to find the best model for measuring the construction of the instrument that has been prepared. The model's suitability is measured by looking at *the standardized root mean square residual* (SRMR) value in the Estimated Model, which is expected < 0.100.

The results of CFA for the teacher performance evaluation model at SDN Doridungga are shown in **Table 3**, where the SRMR value found is 0.001. Since this value is less than 0.100, the model is considered fit. On the other hand, the results of the CFA for the teacher performance evaluation model at SLB Wihdatul Ummah Donggo are shown in **Table 4**, with an SRMR value of 0.258, which means that the model is considered unfit.

Table 3. Model Fit of SDN Doridungga

	Saturated Model	Estimated Model
SRMR	0.001	0.001
d_ULS	0.000	0.000
d_G	N/a	N/a
Chi-Square	Infinite	Infinite
NFI	N/a	N/a

Table 4. Model Fit of SLB Wihdatul Ummah Donggo

	Saturated Model	Estimated Model
SRMR	0.258	0.258
d_ULS	20.018	20.018
d_G	N/a	N/a
Chi-Square	Infinite	Infinite
NFI	N/a	N/a

Furthermore, the t-value (outer loading) and standardized loading factor (SLF) values can be seen in Table 5.

Table 5. CFA Factor Loading Results

Doridungga Elementary School				SLB Wihdatul Ummah Donggo			
Items Code	**Standardized Loading Factor**	**Outer Loading**	**Category**	**Items Code**	**Standardized Loading Factor**	**Outer Loading**	**Category**
I1	0.7	1.00	Valid	I1	0.7	-0.729	Invalid
I2	0.7	0.999	Valid	I2	0.7	-0.074	Invalid
I3	0.7	0.999	Valid	I3	0.7	0.814	Valid
I4	0.7	1.000	Valid	I4	0.7	0.783	Valid
I5	0.7	0.999	Valid	I5	0.7	0.094	Valid
I6	0.7	0.999	Valid	I6	0.7	0.460	Invalid
I7	0.7	0.999	Valid	I7	0.7	0.705	Valid
I8	0.7	0.999	Valid	I8	0.7	0.746	Valid
I9	0.7	0.999	Valid	I9	0.7	0.979	Valid
				I10	0.7	0.912	Valid
				I11	0.7	0.177	Invalid
				I12	0.7	0.859	Valid

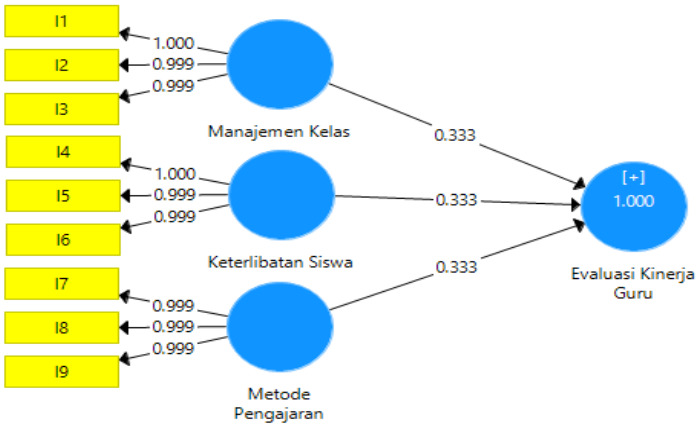

Figure 2. Path Diagram Teacher Performance Model of SDN Doridungga

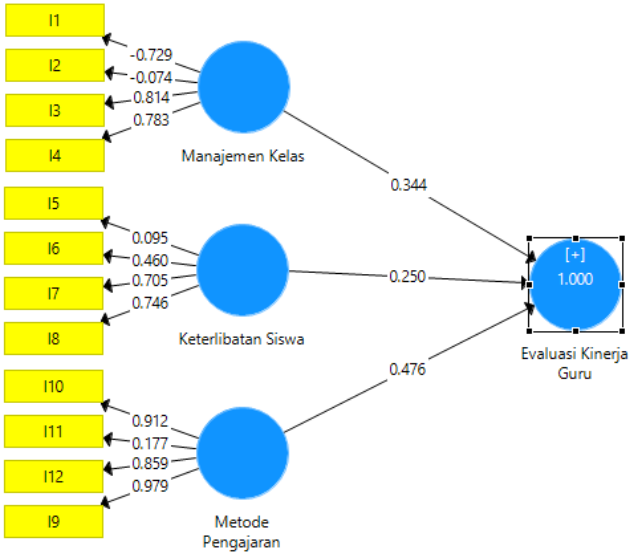

Figure 3. Path Diagram of SLB Teacher Performance Model

In Figures 2 and 3, the outer loading value of each indicator is in each aspect. Meanwhile, the t-value and SLF values can be seen in Table 4, which shows that the item is valid and invalid.

Based on the data in Table 4, all items/variables observed in SDN Doridungga teachers have a significant outer loading value of SLF (> 0.7). Meanwhile, the items/variables observed in SLB Wihdatul Ummah Donggo teachers had a significant average outer loading value from SLF (> 0.7) except for items I1, I2, I6, and I11. This proves that each observed variable used is significant in measuring its latent variables. Overall, it can be said that the observed variables are proven to be reliable and able to measure the construct of teacher performance.

Reliability The results of the calculation of construct reliability using the reliability of the teacher performance assessment instrument are shown in Tables 6 and 7.

Table 6. Construct Reliability of SDN Doridungga

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Teacher Performance Evaluation	**1.000**	**1.000**	**1.000**	**0.998**
Student Engagement	**0.999**	**0.999**	**1.000**	**0.999**
Classroom Management	**0.999**	**0.999**	**1.000**	**0.999**
Teaching Methods	**0.999**	**0.999**	**0.999**	**0.998**

Table 7. Construct Reliability of SLB Wihdatul Ummah Donggo

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Teacher Performance Evaluation	**0.738**	**0.912**	**0.796**	**0.400**
Student Engagement	**0.316**	**0.323**	**0.596**	**0.318**
Classroom Management	**0.018**	**0.723**	**0.224**	**0.453**
Teaching Methods	**0.748**	**0.896**	**0.856**	**0.640**

The results in Table 6 show that all constructions in SDN Doridungga have excellent reliability estimates, with Cronbach's Alpha and Composite Reliability values above 0.7, and AVE more than 0.5. Classroom Management has a good reliability estimate because Cronbach's Alpha and Composite Reliability values exceed 0.7, so the validity is good because the AVE value of 0.999 is more than 0.5. Then Student Engagement has a good reliability estimate because the Alpha and Composite Reliability coefficients exceed 0.7, so the validity is good because the AVE value of 0.999 is more than 0.5. Teaching Methods has a good reliability estimate because all reliability measures exceed 0.7, so the validity is good because the AVE value of 0.998 is more than 0.5. So the instrument is classified as reliable (Heale & Twycross, 2015; Polit, 2015; Quaigrain & Arhin, 2017; Watkins, 2017). On the contrary, the results in Table 7 for the SLB Wihdatul Ummah Donggo model show that Student Engagement and Classroom Management have poor reliability estimates, with Cronbach's Alpha and Composite Reliability values below 0.7, and AVE less than 0.5. Classroom Management has poor reliability estimates because Cronbach's Alpha and Composite Reliability values are less than 0.7, so the validity is not good because the AVE value of 0.453 is less than 0.5. Then, Student Engagement has a poor reliability estimate because the Alpha and Composite Reliability coefficients are less than 0.7. Hence, the validity is not good because the AVE value of 0.318 is more than 0.5. However, the Teaching Method has a good reliability estimate because the value of Cronbach's Alpha and Composite Reliability exceeds 0.7, so the validity is good because the AVE value of 0.640 is more than 0.5. Overall, the teacher performance assessment instrument at SDN Doridungga showed consistency in measurement, while the results at SLB Wihdatul Ummah Donggo needed further attention to improve reliability and validity.

Evaluation Stage

Evaluation of the model's impact on teacher performance showed significant improvements in three main aspects: classroom management, student engagement, and teaching methods. The success indicators set indicate that this model is effective in improving the quality of learning in both schools. The feedback received from teachers and administrators shows increased confidence in using this evaluation model, as well as an increased understanding of expected performance.

Discussion

The teacher performance assessment instrument in regular and inclusive schools consisting of 12 indicators is designed to evaluate three main aspects: classroom management, student engagement, and teaching methods. This instrument is crucial for assessing teacher effectiveness in these two contexts, where diverse student needs must be met. In effective classroom management in the implementation of learning, teachers are required to understand and accommodate various student needs, including those with special education needs (SEN). Teachers must be adept at creating an environment that supports learning for all students and social interaction (Hermanto, 2022).

Student involvement is also a major focus, especially in an inclusive classroom. Teachers must be able to recognize and handle the unique learning needs of each student.

Teacher Efficacy Scale for Inclusive Practice (TEIP-SF) emphasized that teachers' self-efficacy plays an important role in creating an inclusive environment that can encourage active student engagement (Sahli Lozano et al., 2023). This shows that the successful implementation of this evaluation model depends on the teacher's ability to facilitate student engagement in diverse learning environments.

Teaching methods in regular and inclusive classrooms must also be flexible to meet diverse learning needs. Classroom-Based Assessment (CBA) identifies instructional decision-making and educational assessment as critical components of effective teaching methods in inclusive settings (Jungjohann & Gebhardt, 2023). This assessment instrument is expected to be able to provide clearer guidance for teachers in adjusting their teaching methods to be more inclusive and effective.

The results of this study show that the teacher performance assessment instrument developed has good validity and reliability, especially in the evaluation model at SDN Doridungga. With a Cronbach's Alpha value of 0.999 for all constructs, as well as an Average Variance Extracted (AVE) value above 0.5, the instrument can be considered highly reliable and valid. This shows good internal consistency (Taber, 2018; Zakariya, 2022). These findings are in line with previous research that states that instruments that have high reliability contribute to more accurate measurements in teacher performance evaluations (Mohd Razali et al., 2025; Broken & Boon, 2023).

The success of the evaluation model at SDN Doridungga demonstrates the importance of developing instruments that consider the unique characteristics of regular schools. The development of evaluation instruments that take into account the unique characteristics of schools is essential for accurately assessing educational programs (Maisaroh et al., 2024).

However, different results were found in the performance evaluation model in SLB Wihdatul Ummah Donggo, where there are Construction Reliability Challenges in the aspect of Student Engagement and Classroom Management, with a value of Cronbach's Alpha under 0.7. This shows that the instrument may not be able to capture the complexity that exists in the context of SLB learning. Some factors that may influence these outcomes include variations in teaching methods, differences in student needs, and the unique environment at SLB (Husada, 2023; Naila Nur 'Azizah et al., 2024). Therefore, the revision and adaptation of the instrument needed to be more in line with the educational context at SLB Wihdatul Ummah Donggo.

In addition, model analysis using Confirmatory Factor Analysis (CFA) shows that the model for SDN Doridungga has a good fit with the value Standardized Root Mean Square Residual (SRMR) 0.001, while the model for SLB Wihdatul Ummah Donggo shows the value of SRMR 0.258, which indicates that the model is not fit. This emphasizes the need for further review of the items in the instruments used in SLB, as well as the evaluation of better measurement methods that reflect the complexity of learning in SLB Wihdatul Ummah Donggo. These findings support the need for interventions tailored to the psychological and educational needs of students with learning difficulties(Catroppa et al., 2023).

Practical Implications

The findings have significant practical implications for teacher evaluation practices. For regular schools like SDN Doridungga, the model proves to be a reliable tool for enhancing teacher effectiveness and improving student outcomes. In contrast, inclusive schools, such as SLB Wihdatul Ummah Donggo, require further adaptation of the instrument to reflect the unique teaching challenges present in these environments. This suggests that a one-size-fits-all approach to teacher evaluation may not be effective, and there is a need for models that account for the specific requirements of inclusive education.

Limitations and Future Research

One key limitation of this study is the uneven distribution of respondents between regular and inclusive schools, with 22 teachers from SDN Doridungga and 12 teachers from SLB Wihdatul Ummah Donggo. This imbalance may affect the generalizability of the findings. Additionally, the instrument did not fully capture the complexity of teaching in inclusive settings, which requires further refinement and validation. Future research should focus on

developing more specialized evaluation tools for inclusive schools, taking into account the diverse needs of students and the specific competencies required for effective teaching in these contexts.

Conclusion

The teacher performance assessment instrument at SDN Doridungga and SLB Wihdatul Ummah Donggo consists of four main aspects: classroom management (4 points), student involvement (4 points), teaching methods (4 points), and teacher performance evaluation (4 points). Competencies in classroom management encompass teachers' ability to plan and implement learning activities, as well as evaluate and develop student potential. The aspects of student involvement focus on essential personality traits of teachers, such as stability, wisdom, patience, and the ability to serve as positive role models for students. Competence in teaching methods highlights teachers' skills in effective communication with students, peers, parents, and the community.

The analysis results showed that this instrument had valid items, with loading factor values ranging from 0.7 to 1.0, which indicates signifying significance. The model fit was strengthened by the SRMR value for SDN Doridungga of 0.001, which indicates that the model is fit. In contrast, the model for SLB yielded an SRMR value of 0.258, suggesting the need for improvement in that context. Both models demonstrated good reliability, with SDN Doridungga achieving a Cronbach's Alpha score of 1.000. However, at SLB Wihdatul Ummah Donggo, while the reliability score for Teaching Methods reached 0.748, the aspects of Student Engagement and Classroom Management indicated a need for further strengthening.

Given the validated and reliable nature of this instrument, it is deemed suitable for assessing teacher performance at SDN Doridungga. However, specific recommendations for practice include:

1. Tailored Professional Development: Schools, particularly SLB, should implement targeted professional development programs that address the unique challenges faced by teachers in inclusive environments. These programs should focus on adaptive teaching methods, effective classroom management strategies, and engagement techniques that cater to diverse student needs.
2. Regular Review and Adaptation of the Evaluation Model: To ensure continued relevance and effectiveness, the evaluation model should undergo regular reviews and adaptations based on feedback from educators and evolving best practices in inclusive education.

For future research, it is recommended to:

1. Expand the Study to Other Contexts: Conduct similar evaluations in different types of inclusive schools across various regions to determine the adaptability and effectiveness of the assessment instrument in diverse educational settings.
2. Investigate the Impact of Training Programs: Examine the effects of professional development and training programs on teachers' abilities to meet the needs of students with varying abilities, assessing whether such initiatives lead to improved student outcomes.

Overall, these steps will not only enhance the teacher evaluation model's effectiveness but also contribute to the broader goal of improving educational quality in both regular and inclusive schools.

Funding Acknowledgment

This research was funded by the Directorate of Research and Community Service, Ministry of Education, Culture, Research, and Technology (DRPTM) under Contract Number 0459/E5/PG.02.00/2024. We would like to extend our sincere gratitude to LPPM STKIP Harapan Bima for their invaluable support and guidance throughout this study. Their assistance and resources have been instrumental in the successful completion of this research.

Reference

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, *15*(2), 214–221. https://doi.org/10.1016/j.sapharm.2018.03.066

Catroppa, C., Morrison, E., Ryan, N. P., Khan, N., Botchway-Commey, E. N., Moe, S., Piazza, C., Corso, K., Rae, G., Bull, C., McIntosh, E., Anderson, V., Crowe, L., Stonier-Kipen, C., & Sood, N. T. (2023). Characterising the Educational Experiences and Mental Health of Children with Pre-Existing Learning Difficulty or Specific Learning Disorder (SLD) during the COVID-19 Pandemic. *COVID*, *3*(9), 1233–1243. https://doi.org/10.3390/covid3090087

Chukwubikem, P. E. I. (2013). Developing Better Teacher Evaluation. *Journal of Educational and Social Research*, *3*(6), 49–58. https://doi.org/10.5901/jesr.2013.v3n6p49

Cohitmingao, E. S. (2024). Inclusive Education Initiatives: A Narrative Expositions of Teachers. *EPRA International Journal of Environmental Economics, Commerce and Educational Management*, 1–9. https://doi.org/10.36713/epra17630

Dewi, D. E. C., Zubaidah, Z., Lubis, E., & Syaputra, E. (2020). Problematics of Implementation of Inclusion Education in the Elementary School of Bengkulu City. *International Journal of Multicultural and Multireligious Understanding*, *7*(7), 1. https://doi.org/10.18415/ijmmu.v7i7.1701

Gómez López, L. F., & Valdés, M. G. (2019). The Evaluation of Teacher Performance in Higher Education. *Propósitos y Representaciones*, *7*(2), 479–515. https://doi.org/10.20511/pyr2019.v7n2.255

Guía, S. P. G. (2012). *Evaluar la docencia: Estudio meta-evaluador del desarrollo del proceso de encuestas al alumnado y su fiabilidad en el Grado en Biología de la Universidad de Sevilla*. Calidad, evaluación y encuestas de la docencia universitaria.

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence Based Nursing*, *18*(3), 66–67. https://doi.org/10.1136/eb-2015-102129

Hermanto, H. (2022). The teacher performance evaluation in learning management in inclusive settings. *Jurnal Prima Edukasia*, *10*(1), 28–36. https://doi.org/10.21831/jpe.v10i1.37511

Husada, I. A. N. S. (2023). Completeness of Facilities and Infrastructure at Swadaya Kendal Special School (SLB). *International Journal of Science and Society*, *5*(4), 754–764. https://doi.org/10.54783/ijsoc.v5i4.846

Ibrahim, K. (2020). Monitoring and Evaluation of Teacher Effectiveness , A Case of Teacher Performance Appraisal and Development Tool in Public Secondary Schools in Nyandarua South Sub- County, Kenya. *Advances in Social Sciences Research Journal*, *7*(1), 320–329.

Jungjohann, J., & Gebhardt, M. (2023). Dimensions of Classroom-Based Assessments in Inclusive Education. *International Journal of Special Education (IJSE)*, *38*(1), 131–144. https://doi.org/10.52291/ijse.2023.38.12

L'Ecuyer, K. M., Subramaniam, D. S., & Reangsing, C. (2020). Development of the Preceptor Self-Assessment Tool and Use of the Content Validity Index. *The Journal of Continuing Education in Nursing*, *51*(10), 469–476. https://doi.org/10.3928/00220124-20200914-09

Maisaroh, S., Hadi, S., & Andrian, D. (2024). Instrument model to evaluate children-friendly school program at elementary school in Indonesia. *International Journal of Evaluation and Research in Education (IJERE)*, *13*(2), 809. https://doi.org/10.11591/ijere.v13i2.25440

Mardia, M., & Mukhtar S, M. (2022). Analisis Tipologi Dan Gaya Kepemimpinan Dalam Lembaga Pendidikan Islam. *Educational Leadership: Jurnal Manajemen Pendidikan*, *1*(2), 208–225. https://doi.org/10.24252/edu.v1i2.26601

Mohd Razali, M. N., A. Hamid, A. H., Alias, B. S., & Mansor, A. N. (2025). Validity and reliability: Instruments of teacher competency in the context of small schools in Peninsular Malaysia. *Journal of Education and Learning (EduLearn)*, *19*(1), 487–494. https://doi.org/10.11591/edulearn.v19i1.21574

Naila Nur 'Azizah, Miftah Aziza, Zulfiham Hadyan Nadhil, & Difa'ul Husna. (2024). Strategi Dan Proses Pembelajaran PAI Di SLB A Yaketunis Yogyakarta. *Khatulistiwa: Jurnal Pendidikan Dan Sosial Humaniora*, *4*(1), 74–81. https://doi.org/10.55606/khatulistiwa.v4i1.2718

Oya, A., & Haryanto, L. (2022). Penerapan Video Recording Task Untuk Meningkatkan Speaking Skill Bahasa Inggris Mahasiswa Pgsd STKIP Harapan Bima. *Jurnal Ilmiah Mandala Education*, *8*(4), 3094–3103. http://dx.doi.org/10.58258/jime.v8i4.4059

Patah, S. A., & Boon, Y. (2023). Reliability of a Best Practice Instrument for Beginning Teacher Professionalism Development in High-Performing Primary Schools in Malaysia: A Rasch Measurement Model. *Asian Journal of Research in Education and Social Sciences*, *5*(3), 252–258. https://doi.org/10.55057/ajress.2023.5.3.26

Perez-Alvarez, M., Priebe, J., & Susanti, D. (2020). *Teacher Accountability and Pay-for-Performance Schemes in (Semi-) Urban Indonesia*. World Bank, Washington, DC. https://doi.org/10.1596/33376

Permana, A. I., & Eliza, D. (2022). Faktor-Faktor yang Mempengaruhi Kinerja Guru Taman Kanak-Kanak. *Jurnal Basicedu*, *6*(3), 5226–5236. https://doi.org/10.31004/basicedu.v6i3.2824

Polit, D. F. (2015). Assessing measurement in health: Beyond reliability and validity. *International Journal of Nursing Studies*, *52*(11), 1746–1753. https://doi.org/10.1016/j.ijnurstu.2015.07.002

Quaigrain, K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education*, *4*(1), 1301013. https://doi.org/10.1080/2331186X.2017.1301013

Sahli Lozano, C., Wüthrich, S., Baumli, N., Sharma, U., Loreman, T., & Forlin, C. (2023). Development and validation of a short form of the Teacher Efficacy for Inclusive Practices Scale (TEIP-SF). *Journal of Research in Special Educational Needs*, *23*(4), 375–388. https://doi.org/10.1111/1471-3802.12607

Saleh, M., Kumaidi, K., Munadi, S., Mardalis, A., & Subando, J. (2024). The Development of Teacher Performance Assessment Instruments in Boarding School. *European Journal of Educational Research*, *12*(1), 55–68. https://doi.org/10.12973/eu-jer.13.1.55

Saljooghi, B., & Salehi, K. (2016). Developing a Teacher Evaluation Model: The Impact of Teachers' Attitude toward the Performance Evaluation System (PES) on Job Satisfaction and Organizational Commitment with the Mediating Role of Teachers' Sense of Efficacy. *International Journal of Medical Research & Health Sciences*, *5*(5), 200–209.

Scheerens, J. (2013). The use of theory in school effectiveness research revisited. *School Effectiveness and School Improvement*, *24*(1), 1–38. https://doi.org/10.1080/09243453.2012.691100

Shrotryia, V. K., & Dhanda, U. (2019). Content Validity of Assessment Instrument for Employee Engagement. *SAGE Open*, *9*(1), 215824401882175. https://doi.org/10.1177/2158244018821751

Stronge, J. H. (2012). Teacher effectiveness performance evaluation system. *Stronge & Associates*.

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tarhan, H., Karaman, A. C., Kemppinen, L., & Aerila, J.-A. (2019). Understanding Teacher Evaluation in Finland: A Professional Development Framework. *Australian Journal of Teacher Education*, *44*(4), 33–50. https://doi.org/10.14221/ajte.2018v44n4.3

Tenorio, Z., & González Ortega, Y. (2024). Validez aparente y confiabilidad del instrumento "la educación virtual y la satisfacción de los estudiantes en los cursos virtuales." *Enfoque*, *35*(31), 57–66. https://doi.org/10.48204/j.enfoque.v35n31.a5253

Thompson, K., & Lesh, J. (2023). Beyond Standard Measures: Crafting an Inclusive Evaluation Tool for Special Education Teachers in Self-Contained Settings. *Bulletin of Educational Studies*, *2*(2), 65–75. https://doi.org/10.61326/bes.v2i2.125

Watkins, M. W. (2017). The reliability of multidimensional neuropsychological measures: From alpha to omega. *The Clinical Neuropsychologist*, *31*(6–7), 1113–1126. https://doi.org/10.1080/13854046.2017.1317364

Xue, R., Chai, H., Yao, L., & Fu, W. (2023). The influence of school inclusive education climate on physical education teachers' inclusive education competency: The mediating role of teachers' agency. *Frontiers in Psychology*, *14*, 1079853. https://doi.org/10.3389/fpsyg.2023.1079853

Zakariya, Y. F. (2022). Cronbach's Alpha in Mathematics Education Research: Its Appropriateness, Overuse, and Alternatives in Estimating Scale Reliability. *Frontiers in Psychology*, *13*, 1074430. https://doi.org/10.3389/fpsyg.2022.1074430

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. *Journal of Caring Sciences*, *4*(2), 165–178. https://doi.org/10.15171/jcs.2015.017