



Pengaruh Validitas Instrumen dan Reliabilitas Instrumen terhadap Kelayakan Tes dan Non Tes dalam Evaluasi Pembelajaran di SMP Negeri 2 Pangaribuan

Cindi Anjuani Gultom¹, Octavianty Sihombing^{2*}, Sovia Elphani Simarmata³,
Yohanna Sitanggang⁴, Muhammad Bukhori Dalimunthe⁵, Saidun Hutasuhut⁶

cindianjuanigultom@gmail.com¹, sihombingoctavianty10@gmail.com^{2*},

soviasimarmata@gmail.com³, yohannasitanggang246@gmail.com⁴,

daliori86@unimed.ac.id⁵, saidun@unimed.ac.id⁶

^{1,2,3,4,5,6}Program Studi Pendidikan Ekonomi

^{1,2,3,4,5,6}Universitas Negeri Medan

Received: 03 04 2026. Revised: 14 04 2026. Accepted: 29 04 2026.

Abstract : This study aims to analyze the effect of instrument validity and instrument reliability on the feasibility of test and non-test assessments in learning evaluation at SMP Negeri 2 Pangaribuan. High-quality learning evaluation is largely determined by the assessment instruments used; therefore, instruments must meet validity and reliability criteria to produce accurate and trustworthy data. This research employed a quantitative approach with an explanatory research design. The population consisted of teachers and students at SMP Negeri 2 Pangaribuan. Data were collected through questionnaires, documentation, and literature review. Data analysis techniques included validity testing, reliability testing, and regression analysis to determine the relationships among variables. The results indicate that instrument validity has a positive and significant effect on the feasibility of test and non-test assessments in learning evaluation. In addition, instrument reliability also shows a positive and significant effect on the feasibility of assessment instruments. Simultaneously, validity and reliability contribute substantially to the feasibility of test and non-test instruments, indicating that instruments with high levels of validity and reliability are more appropriate for use in learning evaluation. The implications of this study suggest that teachers should pay close attention to validity and reliability aspects when developing assessment instruments to ensure that learning evaluation is conducted effectively and produces accurate data.

Keywords : Instrument validity, Instrument reliability, Instrument feasibility, Learning evaluation.

Abstrak : Penelitian ini bertujuan untuk menganalisis pengaruh validitas instrumen dan reliabilitas instrumen terhadap kelayakan tes dan non-tes dalam evaluasi pembelajaran di SMP Negeri 2 Pangaribuan. Evaluasi pembelajaran yang berkualitas sangat ditentukan oleh instrumen penilaian yang digunakan, sehingga diperlukan instrumen yang memenuhi kriteria valid dan reliabel agar menghasilkan data yang akurat dan dapat dipercaya.

Metode penelitian yang digunakan adalah pendekatan kuantitatif dengan jenis penelitian eksplanatori. Populasi dalam penelitian ini adalah guru dan peserta didik di SMP Negeri 2 Pangaribuan, dengan teknik pengumpulan data menggunakan angket, dokumentasi, dan studi literatur. Analisis data dilakukan melalui uji validitas, uji reliabilitas, serta analisis regresi untuk mengetahui pengaruh antar variabel. Hasil penelitian menunjukkan bahwa validitas instrumen berpengaruh positif dan signifikan terhadap kelayakan tes dan non-tes dalam evaluasi pembelajaran. Selain itu, reliabilitas instrumen juga memiliki pengaruh positif dan signifikan terhadap kelayakan instrumen penilaian. Secara simultan, validitas dan reliabilitas instrumen memberikan kontribusi yang kuat terhadap kelayakan tes dan non-tes, sehingga instrumen yang memiliki tingkat validitas dan reliabilitas tinggi cenderung lebih layak digunakan dalam proses evaluasi pembelajaran. Implikasi dari penelitian ini menunjukkan bahwa guru perlu memperhatikan aspek validitas dan reliabilitas dalam menyusun instrumen penilaian agar evaluasi pembelajaran dapat berjalan secara optimal dan menghasilkan data yang akurat.

Kata Kunci : Validitas instrumen, Reliabilitas instrumen, Kelayakan instrumen, Evaluasi pembelajaran.

PENDAHULUAN

Evaluasi pembelajaran merupakan bagian integral dalam proses pendidikan yang berfungsi untuk mengukur tingkat pencapaian kompetensi peserta didik serta sebagai dasar pengambilan keputusan dalam kegiatan pembelajaran. Evaluasi tidak hanya berperan sebagai alat ukur hasil belajar, tetapi juga sebagai sarana untuk memperbaiki proses pembelajaran secara berkelanjutan. Menurut Suharsimi Arikunto (2013), evaluasi adalah kegiatan untuk mengumpulkan informasi tentang bekerja tidaknya sesuatu yang selanjutnya digunakan sebagai dasar dalam pengambilan keputusan. Dengan demikian, evaluasi pembelajaran harus dilakukan secara sistematis dan menggunakan instrumen yang berkualitas. Instrumen evaluasi pembelajaran pada dasarnya terbagi menjadi dua jenis, yaitu tes dan non tes. Instrumen tes digunakan untuk mengukur kemampuan kognitif peserta didik, sedangkan instrumen non tes digunakan untuk menilai aspek afektif dan psikomotorik seperti sikap, minat, dan keterampilan.

Menurut Anas Sudijono (2015), tes merupakan alat atau prosedur yang digunakan dalam rangka pengukuran dan penilaian di bidang pendidikan, sedangkan teknik *non tes* digunakan untuk memperoleh data yang tidak dapat diukur melalui tes, seperti observasi, wawancara, dan angket. Oleh karena itu, kedua jenis instrumen ini harus memenuhi standar kelayakan agar hasil evaluasi yang diperoleh dapat dipercaya. Kelayakan suatu instrumen evaluasi sangat ditentukan oleh tingkat validitas dan reliabilitasnya. Validitas merupakan ukuran yang menunjukkan tingkat keabsahan suatu instrumen. Menurut Sugiyono (2017), instrumen yang valid berarti alat ukur yang digunakan untuk mendapatkan data adalah valid,

sehingga dapat mengukur apa yang seharusnya diukur. Selain itu, L. J. Cronbach (1971) menegaskan bahwa validitas berkaitan dengan ketepatan interpretasi terhadap hasil pengukuran. Dengan demikian, instrumen yang tidak valid akan menghasilkan data yang tidak sesuai dengan kondisi sebenarnya. Selain validitas, reliabilitas juga merupakan aspek penting dalam menentukan kualitas instrumen.

Reliabilitas berkaitan dengan konsistensi suatu alat ukur. Menurut Arikunto (2013), reliabilitas menunjukkan bahwa suatu instrumen cukup dapat dipercaya untuk digunakan sebagai alat pengumpul data. Sementara itu, Nunnally (1994) menyatakan bahwa reliabilitas mengacu pada konsistensi hasil pengukuran ketika pengukuran dilakukan berulang kali dalam kondisi yang relatif sama. Oleh karena itu, instrumen yang reliabel akan menghasilkan data yang stabil dan dapat diandalkan. Sejumlah penelitian terdahulu menunjukkan bahwa validitas dan reliabilitas instrumen memiliki pengaruh yang signifikan terhadap kualitas dan kelayakan evaluasi pembelajaran. Instrumen yang telah diuji validitas dan reliabilitasnya terbukti mampu meningkatkan akurasi hasil penilaian serta mengurangi kesalahan pengukuran. Sebaliknya, instrumen yang tidak memenuhi kedua kriteria tersebut berpotensi menghasilkan data yang bias sehingga dapat memengaruhi kualitas keputusan pendidikan yang diambil.

Pada praktiknya di sekolah, termasuk di SMP 2 Pangaribuan, masih ditemukan berbagai kendala dalam penyusunan dan penggunaan instrumen evaluasi, baik tes maupun *non tes*. Guru sering kali menggunakan instrumen tanpa melalui proses pengujian validitas dan reliabilitas secara sistematis. Hal ini dapat berdampak pada kurang optimalnya fungsi evaluasi sebagai alat ukur yang objektif dan akurat. Oleh karena itu, diperlukan penelitian yang secara khusus mengkaji pengaruh validitas instrumen (X1) dan reliabilitas instrumen (X2) terhadap kelayakan tes dan *non tes* dalam evaluasi pembelajaran. Studi ini mengevaluasi pengaruh validitas dan reliabilitas instrumen terhadap kelayakan tes serta *non-tes* di SMP 2 Pangaribuan secara parsial dan simultan. Penelitian ini bertujuan memberikan kontribusi teoretis sekaligus praktis dalam meningkatkan standar kualitas evaluasi pembelajaran di lingkungan pendidikan.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian *explanatory research* yang bertujuan untuk mengetahui pengaruh validitas instrumen dan reliabilitas instrumen terhadap kelayakan tes dan *non-tes* dalam evaluasi pembelajaran di SMP Negeri 2 Pangaribuan. Desain penelitian yang digunakan adalah desain korelasional dengan analisis regresi linier berganda, karena penelitian ini melibatkan dua variabel bebas yaitu validitas

instrumen (X1) dan reliabilitas instrumen (X2), serta satu variabel terikat yaitu kelayakan tes dan non-tes (Y). Populasi dalam penelitian ini adalah 35 guru di SMP Negeri 2 Pangaribuan yang menggunakan instrumen evaluasi pembelajaran, dengan teknik pengambilan sampel menggunakan total sampling atau purposive sampling sesuai dengan karakteristik responden yang relevan.

Pengumpulan data dilakukan melalui penyebaran angket (kuesioner) yang disusun menggunakan skala Likert untuk mengukur persepsi responden terhadap validitas, reliabilitas, dan kelayakan instrumen evaluasi pembelajaran. Selain itu, teknik dokumentasi digunakan untuk memperoleh data berupa instrumen tes dan *non-tes* yang digunakan guru, serta wawancara dilakukan sebagai data pendukung apabila diperlukan. Sebelum digunakan, instrumen penelitian terlebih dahulu diuji validitasnya menggunakan korelasi *Product Moment* dan diuji reliabilitasnya menggunakan *Cronbach Alpha* untuk memastikan bahwa instrumen tersebut layak digunakan dalam penelitian.

Teknik analisis data dilakukan melalui beberapa tahapan, yaitu uji prasyarat analisis yang meliputi uji normalitas, uji linearitas, dan uji multikolinearitas. Selanjutnya dilakukan analisis regresi linier berganda untuk mengetahui pengaruh antara variabel bebas terhadap variabel terikat. Pengujian hipotesis dilakukan menggunakan uji t untuk mengetahui pengaruh secara parsial, uji F untuk mengetahui pengaruh secara simultan, serta koefisien determinasi (R^2) untuk melihat besarnya kontribusi variabel validitas dan reliabilitas terhadap kelayakan tes dan *non-tes*. Dengan demikian, hasil penelitian ini diharapkan dapat memberikan gambaran yang jelas mengenai pentingnya validitas dan reliabilitas instrumen dalam menentukan kelayakan evaluasi pembelajaran.

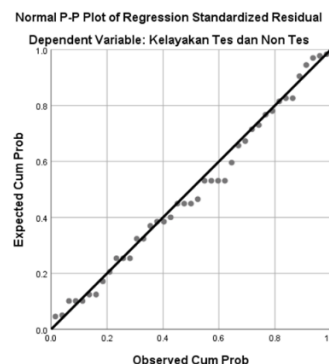
HASIL DAN PEMBAHASAN

Asumsi klasik merupakan serangkaian prasyarat yang harus dipenuhi dalam analisis regresi linier agar model yang dihasilkan bersifat valid, tidak bias, dan dapat digunakan untuk penarikan kesimpulan secara ilmiah. Pengujian asumsi klasik umumnya meliputi uji normalitas, uji multikolinearitas, uji heteroskedastisitas, dan uji autokorelasi. Uji normalitas bertujuan untuk mengetahui apakah data residual berdistribusi normal, sehingga estimasi parameter menjadi lebih akurat. Uji multikolinearitas digunakan untuk mendeteksi adanya korelasi yang tinggi antar variabel independen yang dapat mengganggu kestabilan model regresi. Selanjutnya, uji heteroskedastisitas dilakukan untuk memastikan bahwa varians residual bersifat konstan (homoskedastis), sehingga model tidak mengalami penyimpangan. Sementara

itu, uji autokorelasi bertujuan untuk mengidentifikasi adanya korelasi antar residual pada periode yang berbeda, yang umumnya terjadi pada data *time series*. Dengan terpenuhinya seluruh asumsi klasik tersebut, maka model regresi yang digunakan dalam penelitian dapat dikatakan memenuhi kriteria *Best Linear Unbiased Estimator (BLUE)*, sehingga hasil analisis yang diperoleh lebih dapat dipercaya dan memiliki tingkat akurasi yang tinggi dalam menjelaskan hubungan antar variabel penelitian.

Uji normalitas merupakan salah satu tahapan dalam pengujian asumsi klasik yang bertujuan untuk mengetahui apakah data residual berdistribusi normal atau tidak. Dalam pengambilan keputusan uji normalitas, digunakan nilai signifikansi sebesar 0,05 sebagai batas acuan. Apabila nilai signifikansi lebih besar dari 0,05, maka data dinyatakan berdistribusi normal. Sebaliknya, jika nilai signifikansi lebih kecil dari 0,05, maka data tidak berdistribusi normal. Hasil dari uji normalitas tersebut disajikan pada tabel berikut.

Tabel 1. Hasil Uji Normalitas P-Plot



Uji Normalitas Kolmogorov. Untuk memastikan bahwa residual dalam penelitian ini berdistribusi mendekati normal, dilakukan uji normalitas. Pengujian ini menggunakan metode Kolmogorov-Smirnov untuk mengetahui apakah data residual telah memenuhi asumsi normalitas. Kriteria pengujiannya adalah apabila nilai signifikansi (sig.) lebih besar dari 0,05, maka data dianggap berdistribusi normal. Sebaliknya, jika nilai signifikansi kurang dari 0,05, maka asumsi normalitas tidak terpenuhi. Hasil dari uji normalitas tersebut disajikan pada tabel berikut.

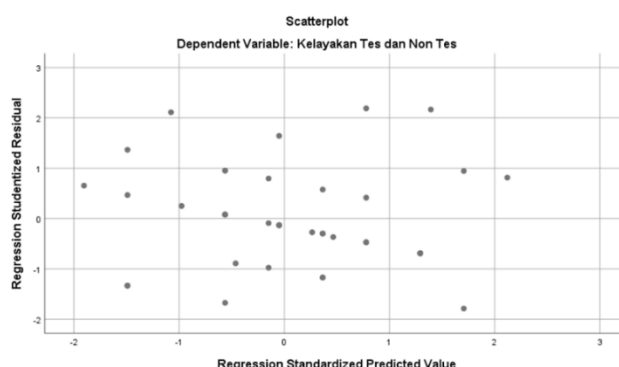
Tabel 2. Hasil Uji Normalitas Kolmogorov-Smirnov

N		41
<i>Normal Parameters^{a,b}</i>	<i>Mean</i>	<i>.0000000</i>
	<i>Std. Deviation</i>	<i>1.13154977</i>
<i>Most Extreme Differences</i>	<i>Absolute</i>	<i>.102</i>
	<i>Positive</i>	<i>.102</i>
	<i>Negative</i>	<i>-.047</i>
<i>Test Statistic</i>		<i>.102</i>

Asymp. Sig. (2-tailed)	.200 ^c
a. Test distribution is Normal	
b. Calculated from data	
c. Lilliefors Significance Correction	
d. This is a lower bound of the significance	

Uji heteroskedastisitas merupakan salah satu tahapan dalam pengujian asumsi klasik yang bertujuan untuk mengidentifikasi apakah terdapat perbedaan varians pada residual dalam model regresi. Model regresi yang baik seharusnya memiliki varians residual yang tetap atau disebut homoskedastisitas. Dalam menentukan hasil pengujian, digunakan nilai signifikansi sebesar 0,05 sebagai dasar pengambilan keputusan. Jika nilai signifikansi lebih besar dari 0,05, maka dapat disimpulkan bahwa model tidak mengalami heteroskedastisitas. Sebaliknya, apabila nilai signifikansi lebih kecil dari 0,05, maka menunjukkan adanya gejala heteroskedastisitas. Oleh karena itu, batas signifikansi 0,05 menjadi acuan penting dalam menilai apakah model regresi telah memenuhi asumsi kesamaan varians.

Tabel 3. Hasil Uji Heteroskedastisitas



Berdasarkan hasil pengujian pada tabel 3 di atas, analisis heteroskedastisitas melalui *scatterplot* menunjukkan bahwa sebaran titik-titik data bersifat acak dan tidak membentuk pola tertentu. Titik-titik tersebut juga tersebar baik di atas maupun di bawah angka nol pada sumbu Y. Dengan demikian, dapat disimpulkan bahwa model regresi yang digunakan tidak mengalami masalah heteroskedastisitas.

Uji multikolinieritas merupakan suatu pengujian yang dilakukan untuk mendeteksi adanya hubungan yang kuat antar variabel independen dalam model regresi. Adanya multikolinieritas dapat menimbulkan ketidakpastian dalam penaksiran koefisien regresi, meningkatkan nilai standar *error*, serta menurunkan ketepatan model dalam melakukan prediksi, sehingga hasil analisis menjadi kurang reliabel. Oleh karena itu, uji ini bertujuan untuk mengetahui apakah terdapat hubungan antar variabel bebas dalam model regresi (Hutasuhut et al., 2024). Salah satu cara yang umum digunakan dalam menguji multikolinieritas adalah

dengan melihat nilai *Variance Inflation Factor* (VIF), di mana nilai $VIF \geq 10$ mengindikasikan adanya multikolinieritas yang serius. Selain itu, nilai toleransi juga digunakan sebagai indikator, di mana nilai toleransi yang rendah menunjukkan tingkat multikolinieritas yang tinggi.

Tabel 4. Hasil Uji Multikolinieritas

Model		Unstandrized Coefficients		Satndardized Coefficients		Collinierity Statistics	
		B	Std. Error	B	t	Sig.	Tolerance VIF
1	(Constant)	10.385	5.633		1.844	.073	
	Validitas Instrumen	.192	.195	.149	1.981	.033	.983 1.017
	Reliabilitas Instrumen	.430	.177	.369	2.434	.020	.983 1.017

a. Dependent Variable: Kelayakan Tes dan Non Tes

Berdasarkan table di atas diperoleh nilai *tolerance value* Validitas Instrumen (X1) dan Reliabilitas Instrumen (X2) lebih besar dari 0,10 ($>0,10$) yaitu sebesar 0,983 begitu juga dengan nilai VIF kedua variabel independent kurang dari 10, yaitu masing-masing bernilai sebesar 1,017, maka dapat disimpulkan model regresi penelitian yang dilakukan tidak terjadi multikolinieritas diantara variabel independent. Untuk menganalisis pengaruh secara simultan antara Gini Ratio yang mencerminkan ketimpangan distribusi pendapatan dan IPM sebagai indikator Indeks Pembangunan Manusia terhadap tingkat kemiskinan di Sumatera Utara selama periode 2012–2025, digunakan metode regresi linier berganda. Regresi linier berganda merupakan pengembangan dari analisis regresi yang digunakan ketika terdapat lebih dari satu variabel independen dalam memprediksi variabel dependen (Rahayu, S., & Hastuti, 2023).

Adapun tujuan penggunaan analisis regresi linier berganda ini adalah untuk mengetahui arah hubungan antar variabel, apakah bersifat positif atau negatif, serta untuk mengestimasi nilai variabel dependen berdasarkan nilai variabel independen yang ada. Hasil analisis ini memberikan informasi yang penting mengenai tingkat signifikansi pengaruh variabel independen terhadap variabel dependen yang diteliti.

Tabel 5. Hasil Analisis Regresi Linier Berganda

Model		Unstandarized Coefficients	
		B	Std. Error
1	(Constant)	10.385	5.633
	Validitas Instrumen	.192	.195
	Reliabilitas Instrumen	.430	.117

a. Dependent Variable: Kelayakan Tes dan Non Tes

Dari hasil uji regresi linier berganda pada table 3 *Coefficients^a* dapat dilihat bahwa diperoleh persamaan dari rumus regresi linier berganda $Y = a + b_1X_1 + b_2X_2$ (Riduwan., 2010) yaitu $Y = 10,385 + 0,192 X_1 + 0,430 X_2$ yang penjelasannya sebagai berikut: 1) (a) merupakan konstanta yang besarnya 10,385 menyatakan bahwa jika variabel independent (Validitas Instrumen dan Reliabilitas Instrumen) sebesar 0 (nol), maka nilai variabel dependent (Kelayakan Tes dan Non Tes) sebesar 10,385. 2) (b_1) merupakan koefisien regresi dari X_1 sebesar 0,192 menyatakan bahwa setiap penambahan satu satuan variabel X_1 dengan asumsi variabel lain (X_2) dianggap konstan maka hal ini akan berpengaruh peningkatan besarnya Y sebesar 0,192. C) (b_2) merupakan koefisin regresi X_2 sebesar 0,430 menyatakan bahwa setiap penambahan satu satuan variabel X_2 dengan asumsi variabel lain (X_1) konstan maka hal ini akan berpengaruh peningkatan besarnya Y sebesar 0,430.

Pengujian hipotesis dilakukan untuk menilai apakah hipotesis yang telah dirumuskan dalam penelitian dapat diterima atau harus ditolak. Dalam konteks penelitian ini, digunakan dua jenis pengujian, yaitu uji t dan uji F, yang bertujuan untuk menganalisis pengaruh variabel bebas Validitas Instrumen (X_1) dan Reliabilitas Instrumen (X_2) terhadap variabel terikat Kelayakan Tes dan Non Tes (Y), baik secara parsial maupun secara simultan. Uji t digunakan untuk menganalisis pengaruh masing-masing variabel independen terhadap variabel dependen secara parsial. Pengujian ini bertujuan untuk mengetahui apakah Validitas Instrumen dan Reliabilitas Instrumen berpengaruh signifikan atau tidak signifikan terhadap Kelayakan Tes dan Non Tes dalam Evaluasi Pembelajaran di SMP Negeri 2 Pangaribuan. Dalam penelitian ini, pengujian dilakukan dengan membandingkan nilai (t_{hitung}) dan (t_{tabel}) pada tingkat signifikansi sebesar 5%. Berdasarkan hasil pengolahan data menggunakan program SPSS versi 25, diperoleh hasil uji t yang disajikan pada tabel berikut.

Tabel 6. Hasil Uji T (Parsial)

Model		<i>Standarized Coefficients</i>		
		<i>Beta</i>	<i>t</i>	<i>Sig.</i>
1	(Constant)		1.844	.073
	Validitas Instrumen	.149	1.981	.033
	Reliabilitas Instrumen	.369	2.434	.020
a. Dependent Variable: Kelayakan Tes dan Non Tes				

Berdasarkan table di atas, diketahui bahwa variabel Validitas Instrumen (X_1) memiliki t_{hitung} sebesar 1,981 dan nilai t_{tabel} diperoleh sebesar 1,68488 dan nilai signifikansi 0,033 yang berarti nilai $t_{hitung} > t_{tabel}$. Maka dapat disimpulkan bahwa hipotesis nol atau H_0 ditolak dan hipotesis alternatif atau H_a diterima, berarti terdapat pengaruh positif dan signifikan antara

variabel Validitas Instrumen (X1) terhadap Kelayakan Tes dan Non Tes (Y) dalam Evaluasi Pembelajaran di SMP Negeri 2 Pangaribuan. Selanjutnya diketahui bahwa pada variabel Reliabilitas Instrumen (X2) memiliki nilai t_{hitung} sebesar 2,434 dan nilai t_{tabel} diperoleh sebesar 1,68488 dan nilai signifikansi 0,020 yang berarti nilai $t_{hitung} > t_{tabel}$. Maka dapat disimpulkan bahwa hipotesis nol atau H_0 ditolak dan hipotesis alternatif H_a diterima, berarti terdapat pengaruh positif dan signifikan antara variabel Reliabilitas Instrumen (X2) terhadap Kelayakan Tes dan Non Tes (Y) dalam Evaluasi Pembelajaran di SMP Negeri 2 Pangaribuan.

Uji F digunakan untuk menganalisis pengaruh variabel independen, yaitu Validitas Instrumen (X1) dan Reliabilitas Instrumen (X2), secara simultan terhadap variabel dependen, yaitu Kelayakan Tes dan Non Tes (Y). Dalam penelitian ini, pengujian dilakukan dengan membandingkan nilai F_{hitung} dengan F_{tabel} , dengan derajat kebebasan pembilang sebesar $(m-1)$ dan penyebut sebesar $(N-m)$, di mana N merupakan jumlah sampel dan m adalah jumlah variabel independen. Berdasarkan hasil pengolahan data menggunakan program SPSS versi 25, diperoleh hasil uji F yang selanjutnya disajikan pada tabel berikut.

Tabel 7. Hasil Uji F (Simultan)

ANOVA ^a						
Model		Sum Of Squares	df	Mean Squares	F	Sig.
1	Regression	8.589	2	4.294	3.186	.000 ^b
	Residual	51.216	38	1.348		
	Total	59.805	40			

a. Dependent Variable: Kelayakan Tes dan Non Tes

b. Predictors: (Constant), Reliabilitas Instrumen, Validitas Instrumen

Berdasarkan table di atas diketahui nilai F_{hitung} pada kolom F yaitu sebesar 3,186 dan nilai F_{tabel} diperoleh sebesar 3,245. Jadi dapat diketahui nilai $F_{hitung} > F_{tabel}$ yaitu $3,186 > 3,245$ dan nilai signifikansi $0,000 < 0,05$, sehingga hipotesis nol atau H_0 ditolak dan hipotesis alternatif H_a diterima. Maka seluruh variabel independen (Validitas Instrumen dan Reliabilitas Instrumen) berpengaruh signifikan secara bersamaan terhadap variabel dependen Kelayakan Tes dan Non Tes dalam Evaluasi Pembelajaran di SMP Negeri 2 Pangaribuan.

Tabel 8. Hasil Uji Koefisien Determinasi

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.379 ^a	.144	.739	1.16095

a. Predictors: (Constant), Validitas Instrumen, Reliabilitas Instrumen
b. Dependent Variable: Kelayakan Tes dan Non Tes

Berdasarkan hasil analisis pada tabel di atas, dapat diketahui bahwa koefisien korelasi pada kolom R sebesar 0,379 dan nilai determinan pada kolom adjusted R Square sebesar 0,739

atau sebesar 73,9%. Hal ini menunjukkan bahwa 73,9% perubahan pada variabel Y yang dipengaruhi oleh perubahan variabel X1 (Validitas Instrumen) dan variabel X2 (Reliabilitas Instrumen). Sedangkan sisanya sebesar 26,1 % dipengaruhi oleh faktor lain dari luar variabel Validitas Instrumen dan variabel Reliabilitas Instrumen yang mempengaruhi variabel Kelayakan Tes dan Non Tes dalam Evaluasi Pembelajaran di SMP Negeri 2 Pangaribuan.

Validitas dan reliabilitas instrumen merupakan determinan utama dalam menilai kelayakan tes dan non tes pada evaluasi pembelajaran di SMP Negeri 2 Pangaribuan. Berdasarkan hasil analisis regresi linier berganda, kedua variabel independen tersebut terbukti memiliki pengaruh positif dan signifikan terhadap variabel dependen, yaitu kelayakan instrumen evaluasi. Secara substantif, temuan ini menunjukkan bahwa peningkatan kualitas validitas instrumen akan berimplikasi langsung pada meningkatnya ketepatan alat ukur dalam merepresentasikan konstruk yang diukur, sedangkan reliabilitas instrumen akan memastikan konsistensi hasil pengukuran yang diperoleh. Dengan demikian, kombinasi antara validitas dan reliabilitas menjadi prasyarat esensial dalam menghasilkan instrumen evaluasi yang tidak hanya akurat, tetapi juga konsisten dan dapat dipertanggungjawabkan secara ilmiah. Pada perspektif teoretis, validitas instrumen memiliki kedudukan yang sangat sentral dalam penelitian pendidikan. (Sugiyono, 2019) menyatakan bahwa validitas merupakan derajat ketepatan antara data yang dilaporkan oleh peneliti dengan data yang sesungguhnya terjadi pada objek penelitian. Artinya, instrumen yang valid mampu mengukur secara tepat apa yang seharusnya diukur, sehingga hasil yang diperoleh benar-benar mencerminkan kondisi empiris di lapangan. (Arikunto, 2018) menegaskan bahwa validitas merupakan syarat mutlak bagi suatu instrumen agar dapat digunakan dalam penelitian ilmiah. Validitas tidak hanya terbatas pada validitas isi, tetapi juga mencakup validitas konstruk dan validitas kriteria, yang secara keseluruhan menentukan sejauh mana instrumen mampu mengungkap fenomena yang diteliti secara komprehensif dan mendalam.

Sementara itu, reliabilitas instrumen berkaitan erat dengan tingkat konsistensi atau keajegan hasil pengukuran. Menurut . (Arikunto, 2018), reliabilitas menunjukkan sejauh mana suatu instrumen dapat memberikan hasil yang tetap apabila digunakan dalam pengukuran yang berulang pada kondisi yang relatif sama. Pendapat ini diperkuat oleh (Sudjana, 2017) yang menyatakan bahwa reliabilitas merupakan indikator penting dalam menjamin kestabilan data yang dihasilkan, sehingga hasil evaluasi dapat dijadikan dasar yang kuat dalam pengambilan keputusan pendidikan. Selain itu, (Anne Anastasi, & Urbina, 1997) mengemukakan bahwa reliabilitas merupakan prasyarat bagi validitas, karena suatu instrumen tidak mungkin memiliki

validitas yang tinggi apabila tidak didukung oleh tingkat konsistensi yang memadai. Hasil uji parsial (uji t) dalam penelitian ini menunjukkan bahwa variabel validitas instrumen dan reliabilitas instrumen masing-masing memiliki pengaruh yang signifikan terhadap kelayakan tes dan *non tes*, yang ditunjukkan oleh nilai signifikansi yang lebih kecil dari 0,05. Hal ini mengindikasikan bahwa setiap peningkatan pada kualitas validitas maupun reliabilitas instrumen akan memberikan kontribusi nyata terhadap peningkatan kualitas instrumen evaluasi pembelajaran. Secara lebih spesifik, validitas instrumen berkontribusi dalam memastikan kesesuaian antara indikator yang diukur dengan tujuan pembelajaran, sedangkan reliabilitas berperan dalam menjaga konsistensi hasil evaluasi sehingga dapat digunakan sebagai dasar penilaian yang objektif.

Hasil uji simultan (uji F) memperlihatkan bahwa validitas dan reliabilitas instrumen secara bersama-sama memberikan pengaruh yang signifikan terhadap kelayakan tes dan *non tes*. Temuan ini mengindikasikan bahwa kedua variabel tersebut memiliki hubungan yang sinergis dalam membentuk kualitas instrumen evaluasi. Dengan kata lain, keberadaan validitas tanpa didukung oleh reliabilitas, ataupun sebaliknya, tidak akan mampu menghasilkan instrumen yang layak digunakan dalam evaluasi pembelajaran. Oleh karena itu, kedua aspek tersebut harus dipenuhi secara simultan agar instrumen yang digunakan benar-benar memenuhi standar kualitas yang diharapkan. Hasil analisis koefisien determinasi menunjukkan bahwa sebagian besar variasi dalam kelayakan tes dan *non tes* dapat dijelaskan oleh variabel validitas dan reliabilitas instrumen, sementara sisanya dipengaruhi oleh faktor lain di luar model penelitian. Faktor-faktor tersebut meliputi karakteristik butir soal, seperti tingkat kesukaran, daya pembeda, efektivitas pengecoh (*distractor*), serta aspek teknis lainnya seperti kejelasan bahasa dan kesesuaian dengan indikator pembelajaran. Hal ini sejalan dengan pandangan (Robert L. Ebel, & Frisbie, 1997) yang menyatakan bahwa kualitas instrumen evaluasi tidak hanya ditentukan oleh validitas dan reliabilitas, tetapi juga oleh kualitas konstruksi butir soal dan kesesuaiannya dengan tujuan pengukuran.

Dari sudut pandang praktis, temuan penelitian ini memberikan implikasi penting bagi guru dan praktisi pendidikan dalam menyusun instrumen evaluasi pembelajaran. Proses penyusunan instrumen sebaiknya tidak hanya berfokus pada penyusunan soal, tetapi juga harus melalui tahapan uji validitas dan reliabilitas secara sistematis, baik melalui uji empiris maupun analisis statistik. Selain itu, diperlukan pula kegiatan revisi dan penyempurnaan instrumen secara berkelanjutan agar kualitas instrumen semakin meningkat. Dengan demikian, instrumen yang digunakan dalam evaluasi pembelajaran tidak hanya memenuhi standar akademik, tetapi

juga mampu memberikan informasi yang akurat dan relevan mengenai pencapaian belajar peserta didik. Secara keseluruhan, hasil penelitian ini memperkuat landasan teoretis yang menyatakan bahwa validitas dan reliabilitas merupakan dua pilar utama dalam evaluasi pembelajaran. Keduanya tidak dapat dipisahkan dan harus dipenuhi secara simultan agar instrumen evaluasi dapat dikatakan layak. Oleh karena itu, peningkatan kualitas instrumen melalui pengujian validitas dan reliabilitas merupakan langkah strategis dalam upaya meningkatkan mutu pendidikan, khususnya dalam menghasilkan sistem evaluasi pembelajaran yang lebih objektif, akurat, dan berorientasi pada peningkatan kualitas hasil belajar peserta didik secara berkelanjutan.

SIMPULAN

Berdasarkan hasil penelitian dan pembahasan, dapat disimpulkan bahwa validitas instrumen memiliki pengaruh yang signifikan terhadap kelayakan tes dan *non-tes* dalam evaluasi pembelajaran. Instrumen yang memiliki tingkat validitas tinggi mampu mengukur kompetensi peserta didik secara tepat sesuai dengan tujuan pembelajaran, sehingga menghasilkan data yang akurat dan dapat digunakan sebagai dasar dalam pengambilan keputusan pendidikan. Dengan demikian, validitas menjadi salah satu faktor utama yang menentukan kualitas suatu instrumen penilaian. Reliabilitas instrumen juga terbukti berpengaruh signifikan terhadap kelayakan tes dan *non-tes*. Instrumen yang reliabel akan menghasilkan data yang konsisten dan stabil meskipun digunakan dalam kondisi yang sama secara berulang. Hal ini menunjukkan bahwa reliabilitas berperan penting dalam menjamin kepercayaan terhadap hasil evaluasi pembelajaran. Instrumen yang tidak reliabel cenderung menghasilkan data yang tidak konsisten, sehingga mengurangi kualitas dan akurasi hasil penilaian. Secara simultan, validitas dan reliabilitas instrumen memiliki pengaruh yang kuat terhadap kelayakan instrumen tes dan *non-tes* dalam evaluasi pembelajaran. Semakin tinggi tingkat validitas dan reliabilitas suatu instrumen, maka semakin tinggi pula tingkat kelayakannya untuk digunakan dalam proses evaluasi pembelajaran.

DAFTAR RUJUKAN

- Aditya, C., & Rahmawati, F. (2025). Pengembangan Pengembangan Instrumen Penilaian Non Tes pada Pembelajaran Berbasis Proyek. *Jurnal Kajian Ilmu Pendidikan (JKIP)*, 6(2), 583-591. <https://doi.org/10.55583/jkip.v6i2.1427>
- Anne Anastasi, & Urbina, S. (1997). *Psychological Testing*. New Jersey.

- Aprianiwati, R., Susanti, T., & Nuraida, N. (2020). Instrumen Asesmen bagi Peserta Didik dalam Proses Pembelajaran Biologi berbasis PBL (Problem Based Learning). *Jurnal BIOEDUIN Program Studi Pendidikan Biologi* 10(2).
<https://doi.org/10.15575/bioeduin.v10i2.12089>.
- Arikunto, S. (2013). *Dasar-dasar evaluasi pendidikan* (Edisi revisi). Jakarta: Bumi Aksara.
- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, S. (2018). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Asrul, A., Saragih, A., & Mukhtar. (2022). *Evaluasi Pembelajaran*. Medan: Perdana Publishing.
- Cronbach, L. J. (1971). Test validation. Dalam R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.
- Hermanto, F. Y., dkk. (2022). Validitas instrumen pembelajaran. *Jurnal Profit*.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Putriyani, & Mutmainnah. (2021). Pengembangan Instrumen Penilaian Autentik. *Jurnal Edumaspul*. <https://jurnal.unimen.cloud/maspuljr/article/view/20/18>
- Riduwan. (2010). *Dasar-Dasar Statistika*. Bandung: Alfabeta