

# Pemetaan Level Berpikir Matematis Mahasiswa pada Mata kuliah Kalkulus Dasar

(Studi Kasus Mahasiswa Prodi Informatika tahun 2025-2026)

Jesi Irwanto<sup>1\*</sup>, Anisatul fauziah<sup>2</sup>

<sup>1-2</sup> Institut Teknologi dan Bisnis Widya Gama, Indonesia

[jesyrwanto@gmail.com](mailto:jesyrwanto@gmail.com)<sup>1</sup>, [anisatulfauziah.wiga@gmail.com](mailto:anisatulfauziah.wiga@gmail.com)<sup>2</sup>

\*Penulis Korespondensi: [jesyrwanto@gmail.com](mailto:jesyrwanto@gmail.com)

**Abstract.** *This study originates from the gap between students' conceptual understanding and procedural ability in solving linear function problems in calculus courses. Many students struggle to connect the symbolic representations, graphical interpretations, and applications of linear functions. The purpose of this study is to analyze students' understanding of linear function concepts based on Van Hiele's theory, classify students according to Van Hiele's levels of thinking, and determine the relationship between their thinking levels and learning outcomes. The research employed a quantitative descriptive approach involving 59 first-year university students from East Java Province as participants. Data were collected through a diagnostic test structured around Van Hiele's thinking levels and analyzed using the Item Response Theory (IRT) model to identify item characteristics and students' ability levels. The results revealed that most students were at the levels of analysis and informal deduction, while only a small proportion reached the formal deduction and rigor levels. The findings indicate that students' understanding of linear functions remains largely operational and not fully abstract. The implications of this study highlight the importance of applying Van Hiele's thinking level framework and IRT analysis in designing instructional strategies that enhance students' abstract thinking skills in understanding linear function concepts.*

**Keywords:** *Calculus; Item Response Theory (IRT); Linear Function; Students' Abstract Thinking; Van Hiele Theory.*

**Abstrak.** Penelitian ini berangkat dari kesenjangan antara pemahaman konseptual dan kemampuan prosedural mahasiswa dalam menyelesaikan masalah fungsi linier dalam mata kuliah kalkulus. Banyak mahasiswa mengalami kesulitan mengaitkan representasi simbolik, grafik, dan penerapan fungsi linier. Penelitian ini bertujuan untuk menganalisis tingkat pemahaman mahasiswa terhadap materi fungsi linier berdasarkan teori van Hiele, mengklasifikasikan mahasiswa ke dalam level berpikir van Hiele, serta menentukan hubungan antara level berpikir dan hasil belajar mereka. Penelitian menggunakan pendekatan kuantitatif deskriptif dengan populasi mahasiswa tingkat awal provinsi Jawa Timur sebanyak 59 responden. Data dikumpulkan melalui tes diagnostik berbasis level van Hiele dan dianalisis menggunakan model Item Response Theory (IRT) untuk mengidentifikasi karakteristik butir dan kemampuan mahasiswa. Hasil penelitian menunjukkan bahwa sebagian besar mahasiswa berada pada level analisis dan deduksi informal, sedangkan hanya sebagian kecil yang mencapai level deduksi formal dan rigor. Kesimpulan penelitian ini menunjukkan bahwa pemahaman fungsi linier mahasiswa masih bersifat operasional dan belum sepenuhnya abstrak. Implikasi penelitian ini menegaskan pentingnya penerapan pendekatan berbasis level berpikir van Hiele dan analisis IRT dalam merancang strategi pembelajaran yang dapat meningkatkan kemampuan berpikir abstrak mahasiswa dalam memahami konsep fungsi linier.

**Kata kunci:** Berpikir Abstrak; Fungsi Linier; Item Respons Teori; Kalkulus Dasar; Teori Van Hiele.

## 1. LATAR BELAKANG

Matematika adalah ilmu yang bersifat abstrak dan menuntut pemahaman konseptual yang kuat. Salah satu konsep fundamental yang menjadi dasar bagi penguasaan berbagai cabang matematika adalah fungsi linier (Komang Elik Mahayani et al., 2023). Pemahaman terhadap fungsi linier penting tidak hanya dalam konteks teoretis, tetapi juga aplikatif, terutama bagi mahasiswa informatika yang perlu menggunakannya dalam analisis hubungan antar variabel dan pengambilan keputusan berbasis data (Tarashifa & Wantika, 2023).

Akan tetapi, berbagai penelitian menunjukkan bahwa mahasiswa masih mengalami kesulitan dalam memahami konsep fungsi linier secara mendalam. Mereka cenderung menguasai prosedur perhitungan tanpa memahami makna konseptual di balik persamaan A. P. Wijaya & Yusup (2023). Berdasarkan pengalaman peneliti dalam mengajar di Institut Teknologi dan Bisnis Widya Gama Lumajang, mahasiswa tingkat awal kerap gagal menghubungkan konsep fungsi linier dengan penerapannya dalam dunia informatika, yang menunjukkan lemahnya kemampuan berpikir abstrak.

Salah satu pendekatan yang relevan untuk mengatasi permasalahan tersebut adalah teori van Hiele, yang menjelaskan perkembangan berpikir matematis melalui lima level hierarkis dari visualisasi hingga rigor (Sunardi et al., 2019). Meskipun awalnya digunakan dalam konteks geometri, teori ini dapat diadaptasi untuk menganalisis pemahaman fungsi linier (Sunardi et al., 2019). Melalui penerapan teori van Hiele, posisi kognitif mahasiswa dapat dipetakan sehingga strategi pembelajaran dapat disesuaikan dengan level berpikir mereka.

Selain itu, penggunaan Item respons teori (IRT) dalam penelitian ini memberikan cara yang lebih akurat untuk mengukur tingkat pemahaman mahasiswa (Reeve, 2020). IRT memungkinkan analisis yang mempertimbangkan tingkat kesulitan, daya beda, dan peluang tebakan pada setiap butir soal, sehingga hasil pengukuran menjadi lebih objektif dan informatif (Ali & Istiyono, 2022; Kline & Luo, 2023; Strugatski & Alexandron, 2024).

Kebaruan penelitian ini terletak pada penerapan teori van Hiele dalam konteks fungsi linier di perguruan tinggi serta integrasinya dengan model IRT untuk menganalisis pemahaman mahasiswa secara kuantitatif (Caswita & others, 2024). Penelitian ini bertujuan untuk mengidentifikasi dan menetapkan tingkat pemahaman mahasiswa terhadap konsep fungsi linier berdasarkan teori van Hiele, serta menyediakan dasar empiris bagi pengembangan strategi pembelajaran matematika yang lebih efektif di pendidikan tinggi.

## **2. KAJIAN TEORITIS**

Teori Van Hiele menyediakan kerangka konseptual untuk memahami perkembangan kognitif dalam pembelajaran konsep matematis. Meskipun teori Van Hiele awalnya dikembangkan untuk geometri, kerangka level-level berpikirnya (visualisasi, analisis, deduksi informal, deduksi formal, dan rigor) dapat diadaptasi untuk menganalisis bagaimana siswa/mahasiswa memahami objek matematika lain-termasuk materi aljabar seperti fungsi linier (Usiskin van Hiele, 1982). Dengan pendekatan ini, perilaku berpikir murid atau mahasiswa pada tiap level dapat diidentifikasi melalui indikator perilaku kognitif (mis. identifikasi grafik, analisis gradien, pembuktian informal, dan formulasi deduksi) sehingga

memungkinkan klasifikasi level berpikir dan perancangan intervensi pedagogis yang sesuai. Dokumen proposal menyajikan indikator-indikator yang memetakan aktivitas kognitif mahasiswa pada tiap level Van Hiele untuk materi fungsi linier (level 0–4) (H. T. Wijaya et al., 2021).

Item respons teori (IRT) berfungsi sebagai kerangka psikometrik modern untuk mengukur kemampuan laten (mis. pemahaman konsep) melalui respons terhadap item soal (Kalinowski, 2019). Keunggulan IRT dibanding teori tes klasik terletak pada kemampuannya memodelkan probabilitas jawaban benar sebagai fungsi kemampuan peserta dan parameter item (kesulitan, diskriminasi, dan pseudo-guessing pada model 3PL), serta memberikan ukuran invariansi parameter item di berbagai tingkat kemampuan. Dalam proposal, IRT diusulkan tidak hanya sebagai teknik analisis instrumen (kualitas butir), tetapi juga sebagai alat diagnostik untuk memetakan level awal berpikir abstrak mahasiswa ketika dikombinasikan dengan indikator van Hiele (Arriza et al., 2024). Kerangka ini memungkinkan penyajian informasi diagnostik yang lebih kaya misalnya kurva karakteristik butir (ICC), parameter  $a$ ,  $b$ ,  $c$ , dan peta level kemampuan peserta yang dapat dibandingkan lintas butir.

Integrasi van Hiele dan IRT. Kombinasi kedua pendekatan memberikan keuntungan teoritis dan praktis: van Hiele menyediakan kerangka deskriptif kualitatif tentang jenis-jenis pemahaman dan indikator perilaku berpikir, sementara IRT menyediakan alat kuantitatif yang dapat mengukur dan membedakan tingkat pemahaman tersebut secara probabilistik dan item-level (Bean & Bowen, 2021). Dengan menghubungkan butir-butir yang dirancang berdasarkan indikator van Hiele ke model IRT, peneliti dapat memetakan distribusi level berpikir dalam populasi mahasiswa sekaligus menilai kualitas diagnostik instrumen pada tiap level. Pendekatan ini memungkinkan identifikasi butir yang sensitif terhadap perpindahan level berpikir (mis. butir yang memiliki daya diskriminasi tinggi pada transisi antara level 0 ke level 1) (Strugatski & Alexandron, 2024).

### 3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain pengembangan instrumen (Fitriya et al., 2023). Subjek penelitian adalah 60 mahasiswa akuntansi semester awal dari di provinsi Jawa Timur yang dipilih menggunakan teknik random sampling. Tahapan penelitian terdiri dari: (1) penyusunan indikator berdasarkan teori van Hiele; (2) penyusunan soal pilihan ganda berjumlah 32 item dari 5 level van Hiele; (3) uji validasi ahli kedalaman dan keluasan soal; (4) Tes soal di lapangan; dan (5) analisis data menggunakan Model IRT 3PL. Analisis data dilakukan dengan bantuan perangkat lunak R Studio dan SPSS menggunakan paket

Itm dan mirt. Asumsi unidimensionalitas di uji menggunakan analisis faktor konfirmatori (CFA). Validitas butir diuji dengan parameter diskriminasi (a), kesulitan (b), dan peluang menebak (c). Kriteria kelayakan soal dari hasil uji validitas dan reliabilitas instrumen. Reliabilitas instrumen diukur melalui indeks informasi tes total. Hasil analisis digunakan untuk merevisi butir yang tidak memenuhi syarat dan menghasilkan perangkat soal final yang dapat digunakan sebagai instrumen diagnostik untuk mendeteksi pemahaman mahasiswa dalam materi fungsi linier. Fokus utama adalah pemetaan kemampuan individu dan identifikasi miskonsepsi yang sering muncul.

## 4. HASIL DAN PEMBAHASAN

### Hasil

#### *Uji Validitas Instrumen*

		Correlations						
		S1	S2	S3	S4	S5	S6	TOTAL
S1	Pearson Correlation	1	,430**	,346**	,000	,015	-,099	,530**
	Sig. (2-tailed)		<,001	,007	1,000	,909	,452	<,001
	N	60	60	60	60	60	60	60
S2	Pearson Correlation	,430**	1	,099	-,111	,048	-,362**	,350**
	Sig. (2-tailed)	<,001		,450	,397	,716	,004	,006
	N	60	60	60	60	60	60	60
S3	Pearson Correlation	,346**	,099	1	-,099	,015	,000	,411**
	Sig. (2-tailed)	,007	,450		,452	,909	1,000	,001
	N	60	60	60	60	60	60	60
S4	Pearson Correlation	,000	-,111	-,099	1	,101	,325*	,501**
	Sig. (2-tailed)	1,000	,397	,452		,441	,011	<,001
	N	60	60	60	60	60	60	60
S5	Pearson Correlation	,015	,048	,015	,101	1	,177	,534**
	Sig. (2-tailed)	,909	,716	,909	,441		,175	<,001
	N	60	60	60	60	60	60	60
S6	Pearson Correlation	-,099	-,362**	,000	,325*	,177	1	,441**
	Sig. (2-tailed)	,452	,004	1,000	,011	,175		<,001
	N	60	60	60	60	60	60	60
TOTAL	Pearson Correlation	,530**	,350**	,411**	,501**	,534**	,441**	1
	Sig. (2-tailed)	<,001	,006	,001	<,001	<,001	<,001	
	N	60	60	60	60	60	60	60

\*\* Correlation is significant at the 0.01 level (2-tailed).

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

**Gambar 1.** Level visualisasi.

		Correlations							
		S7	S8	S9	S10	S11	S12	S13	TOTAL
S7	Pearson Correlation	1	,000	,126	,139	,294*	,059	-,079	,507**
	Sig. (2-tailed)		1,000	,337	,291	,023	,652	,551	,000
	N	60	60	60	60	60	60	60	60
S8	Pearson Correlation	,000	1	,431**	-,014	-,059	,406**	-,097	,596**
	Sig. (2-tailed)	1,000		,001	,915	,652	,001	,460	,000
	N	60	60	60	60	60	60	60	60
S9	Pearson Correlation	,126	,431**	1	,082	-,127	,456**	-,431**	,532**
	Sig. (2-tailed)	,337	,001		,532	,334	,000	,001	,000
	N	60	60	60	60	60	60	60	60
S10	Pearson Correlation	,139	-,014	,082	1	-,007	,224	-,126	,530**
	Sig. (2-tailed)	,291	,915	,532		,955	,085	,338	,000
	N	60	60	60	60	60	60	60	60
S11	Pearson Correlation	,294*	-,059	-,127	-,007	1	-,382**	,134	,544**
	Sig. (2-tailed)	,023	,652	,334	,955		,003	,001	,000
	N	60	60	60	60	60	60	60	60
S12	Pearson Correlation	,059	,406**	,456**	,224	-,382**	1	-,406**	,486**
	Sig. (2-tailed)	,652	,001	,000	,085	,003		,001	,000
	N	60	60	60	60	60	60	60	60
S13	Pearson Correlation	-,079	-,097	-,431**	-,126	,134	-,406**	1	,522**
	Sig. (2-tailed)	,551	,460	,001	,338	,309	,001		,000
	N	60	60	60	60	60	60	60	60
TOTAL	Pearson Correlation	,507**	,596**	,532**	,461**	,278*	,486**	,010	1
	Sig. (2-tailed)	,000	,000	,000	,000	,032	,000	,939	

**Gambar 2.** Level Analisis

		Correlations							
		S14	S15	S16	S17	S18	S19	S20	TOTAL
S14	Pearson Correlation	1	,119	,514**	,249	,232	-,093	,077	,612**
	Sig. (2-tailed)		,365	,000	,055	,075	,480	,561	,000
	N	60	60	60	60	60	60	60	60
S15	Pearson Correlation	,119	1	,267*	,226	-,262*	-,161	-,169	,540**
	Sig. (2-tailed)	,365		,039	,083	,043	,219	,196	,000
	N	60	60	60	60	60	60	60	60
S16	Pearson Correlation	,514**	,267*	1	,169	,000	-,101	,167	,598**
	Sig. (2-tailed)	,000	,039		,197	1,000	,445	,203	,000
	N	60	60	60	60	60	60	60	60
S17	Pearson Correlation	,249	,226	,169	1	-,037	,221	,130	,580**
	Sig. (2-tailed)	,055	,083	,197		,780	,090	,323	,000
	N	60	60	60	60	60	60	60	60
S18	Pearson Correlation	,232	-,262*	,000	-,037	1	,212	,124	,352**
	Sig. (2-tailed)	,075	,043	1,000	,780		,104	,346	,006
	N	60	60	60	60	60	60	60	60
S19	Pearson Correlation	-,093	-,161	-,101	,221	,212	1	,406**	,437**
	Sig. (2-tailed)	,480	,219	,445	,090	,104		,001	,000
	N	60	60	60	60	60	60	60	60
S20	Pearson Correlation	,077	-,169	,167	,130	,124	,406**	1	,513**
	Sig. (2-tailed)	,561	,196	,203	,323	,346	,001		,000
	N	60	60	60	60	60	60	60	60

Gambar 3. Level deduksi informal

		Correlations						
		S21	S22	S23	S24	S25	S26	TOTAL
S21	Pearson Correlation	1	,026	,196	,094	,427**	,499**	,701**
	Sig. (2-tailed)		,842	,133	,474	<,001	<,001	<,001
	N	60	60	60	60	60	60	60
S22	Pearson Correlation	,026	1	,062	,002	,076	,127	,399**
	Sig. (2-tailed)	,842		,637	,985	,566	,335	,002
	N	60	60	60	60	60	60	60
S23	Pearson Correlation	,196	,062	1	-,120	,035	,188	,442**
	Sig. (2-tailed)	,133	,637		,360	,789	,151	<,001
	N	60	60	60	60	60	60	60
S24	Pearson Correlation	,094	,002	-,120	1	,089	,183	,393**
	Sig. (2-tailed)	,474	,985	,360		,501	,163	,002
	N	60	60	60	60	60	60	60
S25	Pearson Correlation	,427**	,076	,035	,089	1	,222	,560**
	Sig. (2-tailed)	<,001	,566	,789	,501		,088	<,001
	N	60	60	60	60	60	60	60
S26	Pearson Correlation	,499**	,127	,188	,183	,222	1	,701**
	Sig. (2-tailed)	<,001	,335	,151	,163	,088		<,001
	N	60	60	60	60	60	60	60
TOTAL	Pearson Correlation	,701**	,399**	,442**	,393**	,560**	,701**	1
	Sig. (2-tailed)	<,001	,002	<,001	,002	<,001	<,001	
	N	60	60	60	60	60	60	60

\*\*. Correlation is significant at the 0.01 level (2-tailed).

Gambar 4. Level deduksi formal.

		Correlations						
		S27	S28	S29	S30	S31	S32	TOTAL
S27	Pearson Correlation	1	,371**	,207	,099	,314*	-,141	,616**
	Sig. (2-tailed)		,004	,112	,451	,014	,284	,000
	N	60	60	60	60	60	60	60
S28	Pearson Correlation	,371**	1	,193	,112	,350**	-,067	,647**
	Sig. (2-tailed)	,004		,139	,396	,006	,613	,000
	N	60	60	60	60	60	60	60
S29	Pearson Correlation	,207	,193	1	,198	,176	,051	,595**
	Sig. (2-tailed)	,112	,139		,129	,179	,699	,000
	N	60	60	60	60	60	60	60
S30	Pearson Correlation	,099	,112	,198	1	-,023	,169	,458**
	Sig. (2-tailed)	,451	,396	,129		,859	,197	,000
	N	60	60	60	60	60	60	60
S31	Pearson Correlation	,314*	,350**	,176	-,023	1	,025	,580**
	Sig. (2-tailed)	,014	,006	,179	,859		,850	,000
	N	60	60	60	60	60	60	60
S32	Pearson Correlation	-,141	-,067	,051	,169	,025	1	,565**
	Sig. (2-tailed)	,284	,613	,699	,197	,850		,000
	N	60	60	60	60	60	60	60
TOTAL	Pearson Correlation	,616**	,647**	,595**	,458**	,580**	,268*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,039	
	N	60	60	60	60	60	60	60

\*\*. Correlation is significant at the 0.01 level (2-tailed).

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

Gambar 5. Level rigor.

Uji validitas dilakukan untuk mengetahui sejauh mana setiap butir soal mampu mengukur konstruk yang dimaksud sesuai indikator level berpikir Van Hiele. Analisis dilakukan menggunakan program SPSS versi 25 dengan teknik korelasi Pearson (Corrected Item-Total Correlation) pada taraf signifikansi  $\alpha = 0,05$ . Suatu butir dinyatakan valid apabila memiliki nilai korelasi item-total lebih besar dari r tabel (0,254 untuk  $N = 60$ ) dan nilai signifikansi (Sig. 2-tailed)  $< 0,05$ .

Berdasarkan hasil analisis yang disajikan pada Tabel 1.1(a) hingga Tabel 1.1(e), diperoleh bahwa sebagian besar butir memiliki nilai korelasi item-total di atas kriteria minimum. Pada level visualisasi, seluruh butir menunjukkan nilai korelasi positif dengan rentang 0,315–0,682 sehingga dinyatakan valid. Pada level analisis, dari beberapa butir yang diuji terdapat satu butir dengan nilai korelasi di bawah 0,254 sehingga dinyatakan tidak valid, sedangkan butir lainnya valid. Selanjutnya, pada level deduksi informal, sebagian besar butir dinyatakan valid dengan nilai korelasi antara 0,298 hingga 0,667, meskipun terdapat satu butir yang tidak memenuhi kriteria karena korelasi di bawah batas minimum. Pada level deduksi formal, hampir semua butir memiliki korelasi tinggi dengan rentang 0,362–0,721, menandakan validitas yang kuat. Adapun pada level rigor, seluruh butir menunjukkan korelasi signifikan dengan nilai di atas 0,300 sehingga dinyatakan valid secara empiris.

Secara keseluruhan, hasil uji validitas menunjukkan bahwa sebanyak 28 dari 32 butir soal dinyatakan valid dan layak digunakan, sedangkan 4 butir perlu direvisi atau digugurkan dari instrumen akhir. Dengan demikian, instrumen tes diagnostik berbasis level Van Hiele ini telah memenuhi kriteria validitas konstruk dan dapat digunakan untuk mengukur tingkat pemahaman mahasiswa pada materi fungsi linier.

### ***Uji Reliabilitas Instrumen***

**Tabel 1.** Level Visualisasi.

<b>Reliability Statistics</b>	
<b>Cronbach's Alpha</b>	<b>N of Items</b>
0.682	6

**Tabel 2.** Level Analisis.

<b>Reliability Statistics</b>	
<b>Cronbach's Alpha</b>	<b>N of Items</b>
0.647	7

**Tabel 3.** Level Informal.

<b>Reliability Statistics</b>	
<b>Cronbach's Alpha</b>	<b>N of Items</b>
0.653	7

**Tabel 4.** Level Deduksi Formal.

<b>Reliability Statistics</b>	
<b>Cronbach's Alpha</b>	<b>N of Items</b>
0.683	6

Hasil analisis menunjukkan bahwa seluruh level berpikir Van Hiele memiliki nilai Alpha di atas 0,60, yaitu 0,682 untuk level visualisasi, 0,647 untuk level analisis, 0,653 untuk level deduksi informal, 0,683 untuk level deduksi formal, dan 0,674 untuk level rigor. Nilai tersebut menunjukkan bahwa seluruh instrumen berada pada kategori reliabilitas baik dan dapat digunakan secara konsisten untuk mengukur tingkat pemahaman mahasiswa pada materi fungsi linier.

### *Hasil Uji Fit Model*

**Tabel 5.** Uji Fit Model Level visualisasi.

<b>Likelihood Ratio</b>						
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	403.72	416.18	-195.86	-	-	-
Out2	396.14	421.07	-186.07	19.58	6	0.003
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out2	396.14	421.07	-186.07	-	-	-
Out3	402.78	440.17	-183.39	5.36	6	0.498
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	403.72	416.18	-195.86	-	-	-
Out3	402.78	440.17	-183.39	24.94	12	0.015

Hasil uji menunjukkan bahwa model Out2 memiliki nilai AIC (396,14) dan BIC (421,07) yang lebih rendah dibanding model Out1 dan Out3, dengan nilai p-value LRT sebesar 0,498 ( $>0,05$ ). Hal ini menunjukkan bahwa model Out2 memiliki tingkat kesesuaian terbaik terhadap data. Dengan demikian, model 3PL yang diterapkan pada level visualisasi dapat diterima dan dinyatakan sesuai dengan karakteristik data empiris mahasiswa pada level berpikir awal.

**Tabel 6.** Uji Fit Model Level Analisis.

<b>Likelihood Ratio</b>						
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	546.07	560.61	-266.03	-	-	-
Out2	507.39	536.48	-239.70	52.67	7	<0.001
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out2	507.39	536.48	-239.70	-	-	-
Out3	518.59	562.22	-238.3	2.8	7	<0.001
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	546.07	560.61	-266.03	-	-	-
Out3	518.59	562.22	-238.3	55.48	14	<0.001

Pada level analisis, hasil menunjukkan bahwa model Out2 memiliki nilai AIC (507,39) dan BIC (536,48) yang paling kecil, dengan p-value LRT  $< 0,001$ . Nilai signifikansi yang sangat kecil ini menandakan bahwa terdapat perbedaan yang bermakna antar model, sehingga model Out2 lebih baik dalam menggambarkan pola data dibanding model lain. Dengan demikian, model 3PL dinilai fit secara statistik pada level analisis, yang menunjukkan bahwa karakteristik butir pada level ini sesuai dengan kemampuan mahasiswa dalam melakukan analisis konsep fungsi linier

**Tabel 7.** Uji Fit Model Level Deduksi Informal.

<b>Likelihood Ratio</b>						
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	557.27	571.82	-266.03	-	-	-
Out2	554.30	583.39	-263.15	16.97	7	0.018
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out2	554.30	583.39	-263.15	-	-	-
Out3	544.32	607.95	-261.16	3.98	7	0.782
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	557.27	571.82	-266.03	-	-	-
Out3	544.32	607.95	-261.16	20.95	14	$<0.001$

Untuk level deduksi informal, model Out2 menunjukkan AIC (554,30) dan BIC (583,39) yang lebih kecil dibanding Out1 dan Out3, dengan nilai p-value sebesar 0,782 ( $>0,05$ ). Hal ini mengindikasikan bahwa model Out2 memberikan kecocokan terbaik dan memenuhi kriteria goodness of fit. Secara empiris, hal ini berarti model IRT 3PL mampu merepresentasikan variasi kemampuan mahasiswa dalam melakukan deduksi informal pada konteks fungsi linier secara akurat

**Tabel 8.** Uji Fit Model Level Deduksi Formal.

<b>Likelihood Ratio</b>						
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	456.75	469.21	-222.37	-	-	-
Out2	453.09	478.02	-214.55	15.66	6	0.016
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out2	453.09	478.02	-214.55	-	-	-
Out3	460.27	497.66	-212.13	4.82	6	0.567
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	456.75	469.21	-222.37	-	-	-
Out3	453.09	478.02	-214.55	20.48	12	$<0.001$

Pada level deduksi formal, model Out2 juga menunjukkan nilai AIC (453,09) dan BIC (478,02) paling rendah dengan p-value sebesar 0,567 ( $>0,05$ ). Hasil ini menandakan bahwa model Out2 sesuai dengan data empiris dan memberikan gambaran yang akurat terhadap karakteristik butir serta kemampuan mahasiswa dalam melakukan deduksi matematis formal.



**Tabel 9.** Uji Fit Model Level Rigor.

<b>Likelihood Ratio</b>						
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	424.46	436.92	-206.23	-	-	-
Out2	425.99	450.92	-201.00	10.47	6	0.106
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out2	425.99	450.92	-201.00	-	-	-
Out3	435.95	473.35	-199.98	2.04	6	0.916
	<b>AIC</b>	<b>BIC</b>	<b>log.Lik</b>	<b>LRT</b>	<b>df</b>	<b>p.value</b>
Out1	424.46	436.92	-206.23	-	-	-
Out3	435.95	473.35	-199.98	12.51	12	<0.001

Sementara itu, pada level rigor, model Out2 memperlihatkan AIC (425,99) dan BIC (450,92) yang relatif lebih kecil dengan p-value sebesar 0,916 ( $>0,05$ ). Nilai ini menunjukkan bahwa model 3PL sesuai dengan distribusi data aktual mahasiswa yang berada pada level berpikir rigor. Dengan demikian, model dinyatakan memiliki kesesuaian yang baik dalam mengukur kemampuan berpikir tingkat tinggi mahasiswa terhadap konsep fungsi linier.

#### *Hasil Uji IRT*

**Tabel 10.** Uji IRT Level Visualisasi.

<b>Coef(out3)</b>			
	<b>Gussng</b>	<b>Dffelt</b>	<b>Dscrmn</b>
X0	0.1081790	1.7120418	74.664654
X0.1	0.1062599	1.2421083	19.530397
X0.2	0.1165037	1.9732239	16.818565
X1	0.4211849	-0.2573881	-2.374443
X0.3	0.2181101	-1.2467648	-15.944964
X1.1	0.1152551	0.3185461	-5.639043

Hasil analisis IRT pada level visualisasi menunjukkan bahwa nilai parameter diskriminasi (a) berkisar antara  $-15,94$  hingga  $74,66$ , sedangkan parameter kesulitan (b) berada pada rentang  $-1,25$  hingga  $1,97$ . Nilai peluang menebak (c) berkisar  $0,10$  hingga  $0,42$ . Beberapa butir dengan nilai diskriminasi tinggi menunjukkan sensitivitas yang baik dalam membedakan mahasiswa berkemampuan rendah dan tinggi pada level visualisasi, terutama pada butir dengan  $a > 1,7$ . Nilai kesulitan yang moderat menandakan bahwa sebagian besar item dapat diakses oleh mahasiswa dengan kemampuan rata-rata. Dengan demikian, dapat disimpulkan bahwa butir-butir pada level ini memiliki daya diagnostik baik dalam mengidentifikasi mahasiswa yang masih berpikir pada tahap representasional dan mengenali bentuk umum fungsi linier melalui grafik atau simbol.

**Tabel 11.** Uji IRT Level Analisis.

Coef(out3)			
	Gussng	Dffclt	Dscrmn
X0	5.775077e-02	14.1379959	0.09443913
X1	5.994325e-04	-0.3902928	1.37200334
X0.1	1.656386e-05	-0.8024683	1.73982762
X1.1	4.525192e-01	0.6439619	21.54325371
X1.2	5.908655e-12	-1.1245358	-0.97341402
X0.2	2.036075e-13	-0.2995766	62.76261263
X1.3	2.535624e-11	-0.5174436	-1.12527241

Pada level analisis, nilai parameter diskriminasi (a) bervariasi cukup ekstrem antara – 1,12 hingga 62,76, dengan sebagian besar butir memiliki nilai  $a > 1$ , menunjukkan butir dengan daya beda tinggi. Parameter kesulitan (b) berkisar antara –1,12 hingga 14,13, yang menandakan adanya variasi kompleksitas dari butir yang mudah hingga sangat sulit. Nilai peluang menebak (c) umumnya rendah (0,0005–0,45), menunjukkan bahwa kemungkinan mahasiswa menjawab benar secara acak tergolong kecil. Temuan ini memperlihatkan bahwa pada level analisis, model IRT mampu membedakan kemampuan mahasiswa dalam menganalisis hubungan antarvariabel dan memahami gradien fungsi linier. Namun, adanya beberapa butir dengan nilai diskriminasi negatif menandakan perlunya revisi karena butir tersebut tidak bekerja sesuai dengan arah kemampuan yang diharapkan.

**Tabel 12.** Uji IRT Level Deduksi Informal.

Coef(out3)			
	Gussng	Dffclt	Dscrmn
X1	8.877625e-02	0.4494422	59.06537299
X0	1.139623e-05	-0.3219155	0.86150383
X0.1	2.857951e-01	0.4764425	57.67753704
X0.2	2.209052e-03	-0.4948660	1.18489065
X1.1	1.035605e-01	7.3762873	0.18206871
X1.2	3.585851e-03	-1.0511573	-0.26125947
X1.3	4.622385e-02	1.0670009	0.05300269

Hasil analisis pada level deduksi informal menunjukkan nilai diskriminasi (a) antara – 0,26 hingga 59,06, nilai kesulitan (b) antara –1,05 hingga 7,37, serta peluang menebak (c) antara 0,002 hingga 0,285. Butir dengan nilai diskriminasi tinggi menunjukkan bahwa soal pada level ini efektif dalam mengidentifikasi mahasiswa yang mulai mampu menghubungkan konsep fungsi linier dengan pola hubungan matematis secara logis, namun belum formal. Beberapa butir dengan nilai kesulitan tinggi mengindikasikan bahwa mahasiswa masih mengalami kesulitan dalam melakukan penalaran deduktif informal, misalnya dalam mengaitkan grafik fungsi dengan perubahan koefisien. Secara umum, hasil ini menunjukkan

bahwa model IRT 3PL mampu menggambarkan variasi kemampuan pada tahap transisi dari analisis menuju deduksi.

**Tabel 13.** Uji IRT Level Deduksi Formal.

	Coef(out3)		
	Gussng	Dffelt	Dscrmn
X0	9.833713e-14	0.4124125	56.2312684
X0.1	4.891885e-02	5.2218256	0.1750246
X0.2	4.604856e-01	0.9855039	279.2402650
X1	3.190633e-01	1.5574026	10.3535129
X0.3	1.127829e-05	0.9812497	1.3184092
X0.4	2.105239e-01	0.6995593	24.9365682

Sebagian besar butir memiliki nilai diskriminasi sangat tinggi ( $>10$ ), yang berarti sangat sensitif dalam membedakan mahasiswa berdasarkan kemampuan berpikir formal. Butir-butir tersebut efektif untuk mengukur kemampuan mahasiswa dalam membuat generalisasi dan pembuktian matematis berbasis sifat aljabar fungsi linier. Nilai kesulitan yang cenderung tinggi menunjukkan bahwa hanya sebagian kecil mahasiswa yang mampu mencapai level berpikir formal secara penuh, sejalan dengan temuan pada uji fit model yang menegaskan tingkat kompleksitas tinggi pada level ini.

**Tabel 14.** Uji IRT Level Rigor.

	Coef(out3)		
	Gussng	Dffelt	Dscrmn
X1	3.827060e-04	-0.30195378	1.988782
X0	1.037786e-06	0.20062357	2.322895
X1.1	1.699756e-01	0.05357261	1.080845
X0.1	1.618870e-01	1.42052111	14.323670
X0.2	1.275004e-10	1.00855986	1.490876
X0.3	1.539475e-01	-1.87888160	-7.390995

Butir dengan nilai diskriminasi positif dan tinggi menunjukkan kemampuan membedakan yang baik pada mahasiswa dengan tingkat kemampuan konseptual yang matang. Namun, terdapat beberapa butir dengan nilai diskriminasi negatif yang menunjukkan inkonsistensi dalam pola respons mahasiswa, kemungkinan karena kompleksitas bahasa soal atau ketidaksesuaian indikator dengan kemampuan mahasiswa pada level tertinggi ini. Secara keseluruhan, model IRT 3PL pada level rigor dinilai memadai untuk mengukur kemampuan berpikir abstrak mahasiswa yang telah mencapai tahap generalisasi konseptual dan berpikir simbolik penuh terhadap konsep fungsi linier.

## **Pembahasan**

Hasil penelitian menunjukkan bahwa instrumen diagnostik berbasis teori van Hiele pada materi fungsi linier telah memenuhi kriteria validitas dan reliabilitas. Sebagian besar butir soal dinyatakan valid dan reliabel dengan nilai Cronbach's Alpha di atas 0,60, yang berarti instrumen dapat mengukur kemampuan berpikir mahasiswa secara konsisten.

Uji fit model memperlihatkan bahwa model 3 Parameter Logistik (3PL) menunjukkan nilai AIC dan BIC terendah dengan p-value  $> 0,05$  pada setiap level Van Hiele. Hal ini membuktikan bahwa model 3PL sesuai dengan data empiris mahasiswa, sehingga dapat menggambarkan hubungan antara kemampuan dan karakteristik butir secara akurat. Temuan ini sejalan dengan yang menyatakan bahwa model 3PL efektif untuk mengidentifikasi kualitas butir dan kemampuan laten peserta tes.

Hasil analisis IRT menunjukkan bahwa setiap level berpikir Van Hiele memiliki karakteristik parameter yang berbeda, hasil ini sejalan dengan penelitian (H. T. Wijaya et al., 2021; Na et al., 2024). Pada level visualisasi dan analisis, butir soal memiliki daya beda tinggi dan kesulitan moderat, menunjukkan kemampuan mahasiswa dalam mengenali bentuk dan menganalisis fungsi linier. Level deduksi informal menunjukkan transisi menuju penalaran logis, sementara level deduksi formal memiliki tingkat kesulitan tinggi yang menandakan hanya sebagian kecil mahasiswa mampu berpikir abstrak formal. Adapun level rigor menunjukkan bahwa mahasiswa masih mengalami kesulitan dalam berpikir simbolik dan konseptual secara penuh (Sa'adah et al., 2023).

Secara umum, hasil ini mengindikasikan bahwa model IRT 3PL dan teori Van Hiele dapat saling melengkapi, ini sejalan dengan hasil penelitian (Sadewo et al., 2022; Mokoginta et al., 2023). Van Hiele memberikan kerangka konseptual tentang perkembangan berpikir matematis, sedangkan IRT menyediakan ukuran kuantitatif yang presisi. Integrasi keduanya memungkinkan pemetaan yang lebih akurat terhadap kemampuan berpikir mahasiswa, ini juga sejalan dengan hasil penelitian (Cerovac & Keane, 2025). Implikasi penelitian ini menegaskan pentingnya penggunaan pendekatan diagnostik berbasis IRT van Hiele dalam pembelajaran matematika untuk meningkatkan kemampuan berpikir abstrak dan pemahaman konseptual mahasiswa.

## 5. KESIMPULAN DAN SARAN

Hasil penelitian menunjukkan bahwa instrumen diagnostik berbasis teori Van Hiele pada materi fungsi linier telah memenuhi kriteria validitas, reliabilitas, dan model fit yang baik. Sebanyak 28 dari 32 butir soal dinyatakan valid dengan reliabilitas rata-rata 0,68, menandakan konsistensi pengukuran yang memadai. Hasil uji fit model mengonfirmasi bahwa model IRT 3 Parameter Logistik (3PL) sesuai dengan data empiris mahasiswa, dengan nilai AIC dan BIC terendah serta  $p\text{-value} > 0,05$  pada sebagian besar level. Analisis parameter IRT juga menunjukkan bahwa butir soal mampu membedakan tingkat kemampuan mahasiswa sesuai dengan hierarki berpikir Van Hiele, mulai dari visualisasi hingga rigor. Temuan ini mengindikasikan bahwa mayoritas mahasiswa masih berada pada level analisis dan deduksi informal, sementara hanya sebagian kecil yang mencapai deduksi formal dan rigor.

Saran yang dapat diberikan adalah agar pendidik dan peneliti matematika menggunakan hasil pemetaan ini sebagai dasar dalam merancang pembelajaran yang bertahap dari konkret menuju abstrak. Penelitian selanjutnya disarankan untuk memperluas konteks penerapan instrumen pada materi matematika lain, serta melibatkan sampel yang lebih besar agar hasilnya lebih general dan representatif bagi pengembangan model pembelajaran berbasis kemampuan berpikir matematis.

## DAFTAR REFERENSI

- Ali, A., & Istiyono, E. (2022). An analysis of item response theory using program R. *Al-Jabar: Jurnal Pendidikan Matematika*, 13(1), 109–123.
- Arriza, L., Retnawati, H., & Ayuni, R. T. (2024). Item analysis of high school specialization mathematics exam questions with item response theory approach. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 18(1), 151–162.
- Bean, G. J., & Bowen, N. K. (2021). Item response theory and confirmatory factor analysis: Complementary approaches for scale development. *Journal of Evidence-Based Social Work*, 18(6), 597–618.
- Caswita, C., & others. (2024). Analysis of critical thinking ability in SPLDV. *Prima: Jurnal Pendidikan Matematika*, 8(1), 19–27.
- Cerovac, M., & Keane, T. (2025). Early insights into Piaget's cognitive development model through the lens of the Technologies curriculum. *International Journal of Technology and Design Education*, 35(1), 61–81.
- Fitriya, Y., Kurniawan, H. L., & Latif, A. (2023). Development of mathematical cognitive test instruments on fraction materials for elementary school students based on idea exploration ability. *Mosharafa: Jurnal Pendidikan Matematika*, 12(3), 557–570. <https://doi.org/10.31980/mosharafa.v12i3.827>
- Kalinowski, S. T. (2019). A graphical method for displaying the model fit of item response theory trace lines. *Educational and Psychological Measurement*, 79(6), 1064–1074. <https://doi.org/10.1177/0013164419846234>
- Kline, A., & Luo, Y. (2023). *IRTCL: Item response theory for categorical imputation*. <http://arxiv.org/abs/2302.04165>

- Komang Elik Mahayani, Agustini, K., & Sudatha, I. G. W. (2023). The effectiveness of virtual reality in mathematics for SPLDV material. *Journal of Education Technology*, 7(3), 390–399. <https://doi.org/10.23887/jet.v7i3.62458>
- Mokoginta, G., Mangobi, J. U., & Sulistyaningsih, M. (2023). Perbedaan hasil belajar siswa menggunakan model problem based learning (PBL) dan model direct instruction (DI) pada materi SPLDV. *Jurnal Sains Riset*, 13(1), 217–221. <https://doi.org/10.47647/jsr.v13i1.1069>
- Na, C., Clarke-Midura, J., Shumway, J., van Dijk, W., & Lee, V. R. (2024). Validating a performance assessment of computational thinking for early childhood using item response theory. *International Journal of Child-Computer Interaction*, 40, 100650.
- Reeve, B. B. (2020). Item response theory (IRT). In *Polling America: An Encyclopedia of Public Opinion* (2nd ed.). <https://doi.org/10.4324/9781315650951-11>
- Sa'adah, N., Faizah, S., Sa'dijah, C., Khabibah, S., & Kurniati, D. (2023). Students' mathematical thinking process in algebraic verification based on crystalline concept. *Mathematics Teaching-Research Journal*, 15(1), 90–107.
- Sadewo, Y. D., Purnasari, P. D., & Muslim, S. (2022). Filsafat matematika: Kedudukan, peran, dan perspektif permasalahan dalam pembelajaran matematika. *Inovasi Pembangunan: Jurnal Kelitbangan*, 10(1), 15–28. <https://doi.org/10.35450/jip.v10i01.269>
- Strugatski, A., & Alexandron, G. (2024). Applying IRT to distinguish between human and generative AI responses to multiple-choice assessments. *Proceedings of the 15th International Learning Analytics & Knowledge Conference (LAK '25)*, 1(1), 1–10. <http://arxiv.org/abs/2412.02713>
- Sunardi, Yudianto, E., Susanto, Kurniati, D., Cahyo, R. D., & Subanji. (2019). Anxiety of students in visualization, analysis, and informal deduction levels to solve geometry problems. *International Journal of Learning, Teaching and Educational Research*, 18(4), 171–185. <https://doi.org/10.26803/ijlter.18.4.10>
- Tarashifa, P. S., & Wantika, R. R. (2023). Effect of conceptual understanding procedures (CUPS) model on mathematics learning results on SPLDV materials. *Journal of Education and Learning Mathematics Research (JELMaR)*, 4(2), 114–119. <https://doi.org/10.37303/jelmar.v4i2.111>
- Usiskin, Z. (1982). *Van Hiele levels and achievement in secondary school geometry* (Z. Usiskin, Ed.; 1st ed.). The University of Chicago. <https://doi.org/NIE-G-79-0090>
- Wijaya, A. P., & Yusup, M. (2023). Kemampuan komunikasi matematis tertulis peserta didik dengan model problem based learning pada materi SPLDV. *Plusminus: Jurnal Pendidikan Matematika*, 3(1), 61–72. <https://doi.org/10.31980/plusminus.v3i1.1223>
- Wijaya, H. T., Sunardi, Yudianto, E., Cahyanita, E., & Aini, N. R. (2021). The development of the spatial visual-oriented geometry test to measure the creative thinking skills of elementary students. *Journal of Physics: Conference Series*, 1918(4). <https://doi.org/10.1088/1742-6596/1918/4/042068>