

DISPARITAS KETERAMPILAN BERPIKIR KRITIS MAHASISWA CALON GURU BIOLOGI DITINJAU MELALUI PENDEKATAN RASCH MODEL

DISPARITIES IN CRITICAL THINKING SKILLS AMONG PROSPECTIVE BIOLOGY TEACHERS AS ASSESSED USING THE RASCH MODEL

Sulton Nawawi¹⁾, Asni Johari^{2)*}, Revis Asra³⁾, Evita Anggereini⁴⁾

^{1) 2*) 3) 4)}Doktor Pendidikan MIPA, Pascasarjana, Universitas Jambi, Jambi, Indonesia,
email:¹⁾sulton.bio@gmail.com, ²⁾johari_asni@yahoo.com (email korespondensi), ³⁾revisasra@unja.ac.id,
⁴⁾evita_pklh@yahoo.com

Diterima: Februari 2026; Disetujui: April 2026; Diterbitkan: Agustus 2026

Abstrak

Keterampilan berpikir kritis merupakan kompetensi penting bagi calon guru biologi dalam mendukung kemampuan analisis, evaluasi, dan pemecahan masalah ilmiah. Penelitian ini bertujuan mendeskripsikan profil dan disparitas keterampilan berpikir kritis calon guru biologi menggunakan analisis *Rasch*. Penelitian menggunakan pendekatan kuantitatif deskriptif terhadap 69 mahasiswa dari dua perguruan tinggi di kota Palembang. Instrumen menggunakan tes berpikir kritis yang dikembangkan oleh *Lauren B. Starkey*, lalu dianalisis menggunakan model *Rasch* melalui *Winsteps* versi 3.73, dimana dari hasil penelitian menunjukkan instrumen memiliki reliabilitas aitem 0,88, reliabilitas responden 0,86, *Cronbach's Alpha* 0,89, dan memenuhi kriteria minimal unidimensionalitas (31,2%). Hasil penelitian pada 2 kampus menunjukkan bahwa sebanyak 52,17% berada pada kategori sangat rendah, 4,35 % kategori rendah, 21,74% kategori tinggi, dan 21,74% kategori sangat tinggi. Namun, pada kampus A didominasi kategori sangat rendah dan rendah 60,87% sedangkan kampus B didominasi 52,16% di kategori sangat tinggi dan tinggi. Meski demikian, berdasarkan hasil *Differential Item Functioning* (DIF) ditemukan bahwa kampus A ternyata memiliki kelebihan dalam mengevaluasi informasi, menilai kualitas argumen, serta memberikan pertimbangan terhadap alternatif solusi berdasarkan bukti yang tersedia. Temuan ini menunjukkan bahwa pendekatan *Rasch* dapat mengungkap profil dan disparitas keterampilan berpikir kritis calon guru biologi, sekaligus menjadi dasar penguatan pengembangan keterampilan berpikir kritis calon guru biologi.

Kata kunci: keterampilan berpikir kritis, calon guru biologi, analisis *Rasch*

Abstract

Critical thinking skills are a crucial competency for prospective biology teachers, supporting their ability to analyze, evaluate, and solve scientific problems. This study aims to describe the profile and disparities in critical thinking skills among prospective biology teachers using *Rasch* analysis. A descriptive quantitative approach was employed, involving 69 university students from two institutions in Palembang. The study utilized a critical thinking test developed by *Lauren B. Starkey*, with data analyzed via the *Rasch* model using *Winsteps* version 3.73. Results indicated that the instrument possessed an item reliability of 0.88, a person reliability of 0.86, and a *Cronbach's Alpha* of 0.89, while also meeting the minimum unidimensionality criterion (31.2%). Across both campuses, 52.17% of students fell into the "very low" category, 4.35% into "low," 21.74% into "high," and 21.74% into "very high." However, Campus A was dominated by the "very low" and "low" categories (60.87%), whereas Campus B was dominated by the "very high" and "high" categories (52.16%). Nevertheless, *Differential Item Functioning* (DIF) analysis revealed that Campus A students demonstrated strengths in evaluating information, assessing argument quality, and considering alternative solutions based on available evidence. These findings demonstrate that the *Rasch* approach can effectively reveal the profiles and disparities in critical thinking skills among prospective biology teachers, thereby serving as a foundation for strengthening the development of these skills.

Keywords: critical thinking skills, prospective biology teachers, *Rasch* analysis

Didaktika Biologi: Jurnal Penelitian Pendidikan Biologi
p-ISSN 2549-5267
e-ISSN 2579-7352

Pendahuluan

Keterampilan berpikir kritis merupakan salah satu kompetensi penting pada abad 21 yang perlu dimiliki oleh

seorang guru, termasuk calon guru biologi. Perkembangan ilmu pengetahuan dan teknologi yang semakin pesat menuntut seseorang mampu menganalisis informasi

lebih objektif, menilai suatu pendapat dengan baik, mengambil keputusan secara logis dan memecahkan masalah secara ilmiah. Dalam dunia pendidikan, keterampilan berpikir kritis sangat penting karena guru tidak hanya menyampaikan materi pembelajaran, tetapi juga membimbing peserta didik agar mampu berpikir analitis, reflektif, dan berdasarkan fakta. Oleh karena itu keterampilan berpikir kritis menjadi salah satu kompetensi utama yang harus dimiliki dalam pembelajaran abad ke-21 (García-Carmona, 2025; Marangio et al., 2024; Yuan & Liao, 2023).

Dalam pembelajaran biologi, keterampilan berpikir kritis memiliki peran yang sangat penting karena biologi mempelajari fenomena kehidupan yang kompleks, kontekstual dan terus berkembang. Pembelajaran biologi tidak hanya menekankan penguasaan konsep, tetapi juga kemampuan memahami fenomena ilmiah, menafsirkan data, menganalisis hubungan sebab-akibat, serta mengambil keputusan terhadap isu-isu sosiosaintifik yang berkembang di masyarakat. Oleh karena itu, calon guru biologi perlu memiliki keterampilan berpikir kritis agar mampu menilai informasi ilmiah secara objektif dan membimbing peserta didik menghadapi tantangan global abad ke-21 (Maryuningsih et al., 2020).

Meskipun demikian, berbagai penelitian menunjukkan bahwa keterampilan berpikir kritis calon guru biologi masih belum berkembang secara optimal. Secara umum telah mampu memahami konsep-konsep dasar biologi, tetapi masih mengalami kesulitan ketika dihadapkan pada aktivitas berpikir tingkat tinggi yang menuntut kemampuan menginterpretasikan informasi, mengevaluasi argumen, menarik inferensi, dan mengambil keputusan berdasarkan bukti ilmiah (Apriliani et al., 2022; Syahfitri, 2023). Kondisi tersebut menunjukkan bahwa peningkatan jenjang akademik belum tentu diikuti oleh berkembangnya keterampilan berpikir kritis apabila proses pembelajaran belum secara konsisten memberikan pengalaman belajar yang mendorong penalaran ilmiah.

Kecenderungan tersebut juga tampak pada berbagai konteks pembelajaran biologi. Pada pembelajaran terkait etika genetika, mahasiswa masih mengalami kesulitan dalam

menyusun argumentasi ilmiah dan mengevaluasi alternatif penyelesaian masalah meskipun pembelajaran berbasis diskusi telah diterapkan (Maryuningsih et al., 2020). Demikian pula pada konteks literasi botani, mahasiswa relatif mampu membaca data dan grafik, tetapi belum optimal dalam mengevaluasi kredibilitas sumber informasi serta mengidentifikasi argumen ilmiah yang valid (Anggrella & Sudrajat, 2025). Temuan-temuan tersebut menunjukkan bahwa tantangan utama dalam pendidikan calon guru biologi tidak lagi terbatas pada penguasaan konsep, melainkan juga pada kemampuan mengevaluasi informasi, mengintegrasikan bukti ilmiah, dan membangun penalaran yang logis dalam memecahkan permasalahan biologi.

Berbagai model pembelajaran inovatif telah dikembangkan untuk meningkatkan keterampilan berpikir kritis, seperti penerapan *Project-Based Learning*-RAHMA, pembelajaran berbasis teknologi digital, dan pembelajaran aktif yang berorientasi pada pengembangan keterampilan berpikir tingkat tinggi (Aflalo, 2026; Darussyamsu et al., 2024; Madieva et al., 2025). Meskipun berbagai penelitian melaporkan bahwa pendekatan tersebut berpotensi meningkatkan keterampilan berpikir kritis, sebagian besar penelitian masih berfokus pada pengujian efektivitas model pembelajaran. Sementara itu, penelitian yang memetakan profil kemampuan berpikir kritis mahasiswa, disparitas kemampuan antar perguruan tinggi, serta karakteristik respons mahasiswa berdasarkan analisis *Rasch* masih relatif terbatas. Padahal, informasi tersebut penting untuk memberikan gambaran objektif mengenai tingkat kemampuan berpikir kritis mahasiswa, mengidentifikasi aspek-aspek keterampilan yang masih memerlukan penguatan, serta menjadi dasar dalam merancang strategi pembelajaran yang lebih efektif bagi calon guru biologi.

Untuk mengukur keterampilan berpikir kritis secara objektif, diperlukan instrumen yang valid dan mampu memberikan estimasi kemampuan yang akurat. Selama ini, pengukuran dalam penelitian pendidikan umumnya menggunakan *Classical Test Theory* (CTT) karena prosedur analisisnya relatif sederhana

serta mampu mengevaluasi reliabilitas dan karakteristik butir berdasarkan skor total responden. Namun, Rusch et al., (2017) menjelaskan bahwa CTT memiliki keterbatasan karena estimasi kemampuan responden dan parameter butir sangat bergantung pada karakteristik sampel yang digunakan, sehingga hasil pengukuran kurang stabil ketika instrumen diterapkan pada kelompok yang berbeda. Selain itu, CTT belum mampu memberikan informasi diagnostik mengenai pola respons individu maupun mendeteksi bias butir secara komprehensif.

Sebagai alternatif, Bond et al., (2020) menjelaskan bahwa Model *Rasch*, sebagai bagian dari *Item Response Theory* (IRT), memungkinkan estimasi kemampuan responden dan tingkat kesulitan butir dilakukan secara independen pada satu skala logit. Model ini tidak hanya menghasilkan ukuran yang lebih objektif dan stabil, tetapi juga menyediakan informasi psikometrik yang lebih komprehensif, seperti *unidimensionalitas*, *person fit*, *item fit*, *separation index*, dan *Differential Item Functioning* (DIF) untuk mengidentifikasi potensi bias butir antar kelompok.

Pada pembelajaran biologi analisis menggunakan model *Rasch* dinilai efektif untuk mengevaluasi kualitas soal berpikir kritis dan kreativitas (Sari & Rakhmawati, 2024). Penggunaan model *Rasch* pada instrumen berbasis konteks biologi materi fotosintesis juga memberikan informasi yang lebih lengkap mengenai profil kemampuan peserta didik, sehingga hasil pengukuran dapat digunakan sebagai dasar untuk memperbaiki proses pembelajaran (Suwita et al., 2024).

Meskipun penelitian mengenai keterampilan berpikir kritis dan analisis *Rasch* telah banyak dilakukan, penelitian yang secara khusus membandingkan profil keterampilan berpikir kritis mahasiswa calon guru biologi antar perguruan tinggi masih relatif terbatas. Padahal, variasi karakteristik pembelajaran pada setiap perguruan tinggi berpotensi menghasilkan capaian keterampilan berpikir kritis yang berbeda. Informasi mengenai disparitas tersebut penting untuk mengevaluasi efektivitas proses pembelajaran sekaligus mengidentifikasi praktik pembelajaran yang

lebih optimal dalam mengembangkan keterampilan berpikir kritis mahasiswa calon guru biologi. Oleh karena itu, penelitian ini bertujuan mendeskripsikan profil dan disparitas keterampilan berpikir kritis mahasiswa calon guru biologi pada dua perguruan tinggi menggunakan model *Rasch* sebagai dasar pengembangan pembelajaran dan asesmen yang lebih efektif.

Oleh karena itu, penelitian ini bertujuan untuk (1) mendeskripsikan profil keterampilan berpikir kritis mahasiswa calon guru biologi pada dua perguruan tinggi menggunakan model *Rasch*, (2) menganalisis disparitas kemampuan berpikir kritis antarperguruan tinggi berdasarkan *person measure*, (3) mengidentifikasi pola respons mahasiswa, tingkat kesulitan butir, dan potensi bias butir melalui analisis DIF, serta (4) memberikan informasi empiris sebagai dasar pengembangan pembelajaran dan asesmen yang lebih efektif dalam meningkatkan keterampilan berpikir kritis calon guru biologi.

Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif untuk memetakan profil keterampilan berpikir kritis calon guru biologi menggunakan analisis *Rasch*. Metode deskriptif dipilih karena penelitian ini bertujuan untuk mendeskripsikan kemampuan berpikir kritis mahasiswa secara objektif berdasarkan hasil pengukuran instrumen tanpa memberikan perlakuan tertentu.

Subjek penelitian terdiri atas 69 mahasiswa calon guru biologi yang berasal dari dua perguruan tinggi, yaitu Kampus A sebanyak 46 mahasiswa dan Kampus B sebanyak 23 mahasiswa. Pemilihan subjek dilakukan menggunakan teknik *purposive sampling* dengan mempertimbangkan mahasiswa yang telah menempuh mata kuliah kependidikan dan biologi dasar sehingga dianggap memiliki pengalaman pembelajaran yang relevan dengan keterampilan berpikir kritis.

Instrumen soal yang digunakan dalam penelitian ini adalah Penilaian Berpikir Kritis yang dikembangkan oleh Lauren B. Starkey. Instrumen ini terdiri dari 30 soal pilihan ganda berpikir kritis dengan skor maksimum

100. Data penelitian diperoleh melalui pelaksanaan tes keterampilan berpikir kritis yang diberikan kepada seluruh responden. Jawaban mahasiswa kemudian diberi skor sesuai rubrik penilaian keterampilan berpikir kritis. Hasil skor selanjutnya dianalisis menggunakan pendekatan *Rasch* untuk memperoleh informasi mengenai kemampuan mahasiswa dan karakteristik instrumen.

Analisis data dilakukan menggunakan model *Rasch* berbantuan perangkat lunak *Winsteps* versi 3.73. Seluruh analisis dilakukan berdasarkan data penelitian utama yang melibatkan 69 mahasiswa calon guru biologi, sehingga parameter psikometrik instrumen merepresentasikan karakteristik responden penelitian dan tidak berasal dari data *pilot test*. Analisis diawali dengan mengestimasi *person measure* untuk memperoleh kemampuan berpikir kritis setiap mahasiswa dalam satuan logit serta *item measure* untuk menentukan tingkat kesulitan setiap butir soal. Kategori kemampuan berpikir kritis ditentukan berdasarkan nilai *mean person measure* dan standar deviasi, sedangkan tingkat kesulitan butir ditentukan berdasarkan mean item dan standar deviasi. Selanjutnya dilakukan analisis distribusi kemampuan berpikir kritis secara keseluruhan dan berdasarkan asal perguruan tinggi untuk memperoleh gambaran disparitas profil kemampuan mahasiswa. Analisis tingkat kesulitan butir juga dilakukan berdasarkan indikator berpikir kritis untuk mengidentifikasi aspek kemampuan yang relatif mudah maupun sulit dikuasai mahasiswa.

Sebagai analisis pendukung, dilakukan *Differential Item Functioning* (DIF) untuk mengidentifikasi adanya butir yang memiliki karakteristik berbeda antara mahasiswa Kampus A dan Kampus B. Suatu butir dinyatakan mengalami DIF apabila memiliki nilai probabilitas (*p-value*) kurang dari 0,05. Analisis DIF digunakan untuk melengkapi interpretasi mengenai disparitas kemampuan berpikir kritis antar kelompok mahasiswa serta mengidentifikasi indikator yang menunjukkan perbedaan karakteristik respons menggunakan kerangka berpikir kritis milik Facione. Facione dipilih sebagai

kerangka analisis karena dimensinya memberikan klasifikasi yang jelas dan komprehensif untuk memetakan keterampilan berpikir kritis yang direpresentasikan dalam setiap butir instrumen.

Instrumen dalam penelitian ini memiliki kualitas yang baik dengan *item reliability* sebesar 0,88, *person reliability* sebesar 0,86, dan *Cronbach's Alpha* sebesar 0,89. Nilai *item separation* sebesar 2,72 dan *person separation* sebesar 2,50 menunjukkan bahwa instrumen mampu membedakan tingkat kesulitan butir maupun kemampuan responden secara memadai. Selain itu, *nilai raw variance explained by measures* sebesar 31,2% menunjukkan bahwa instrumen memenuhi asumsi *unidimensionalitas*, sehingga layak digunakan untuk mengukur satu konstruk dominan, yaitu keterampilan berpikir kritis. Hasil analisis juga menunjukkan bahwa fungsi distraktor bekerja dengan baik sehingga setiap opsi jawaban mampu membedakan kemampuan responden.

Hasil dan Pembahasan

Profil Keterampilan Berpikir Kritis Mahasiswa

Analisis kemampuan berpikir kritis mahasiswa dilakukan berdasarkan nilai *person measure* hasil pemodelan *Rasch*. Pengelompokan kategori kemampuan mengacu pada nilai rata-rata (*mean person measure*) dan standar deviasi. Hasil analisis menunjukkan bahwa nilai *mean person measure* mahasiswa sebesar 0,12 logit dengan standar deviasi (SD) sebesar 1,27 logit. Nilai rata-rata tersebut menunjukkan bahwa secara umum kemampuan mahasiswa berada sedikit di atas rata-rata tingkat kesulitan instrumen (0 logit). Namun, nilai standar deviasi yang relatif besar mengindikasikan adanya variasi kemampuan yang cukup tinggi di antara mahasiswa sehingga kemampuan berpikir kritis belum terdistribusi secara merata. Berdasarkan nilai *mean* dan *standar deviasi*, kemampuan berpikir kritis mahasiswa dikelompokkan ke dalam empat kategori sebagaimana disajikan pada Tabel 1.

Tabel 1. Kategori Kemampuan Berpikir Kritis Berdasarkan *Person Measure*

Kategori	Rentang Logit
Sangat Tinggi	> 1,39
Tinggi	0,12 – 1,39
Rendah	-1,15 – 0,11
Sangat Rendah	< -1,15

Keterangan: Mean = 0,12 logit; SD = 1,27 logit.

Tabel 2. Distribusi Kemampuan Berpikir Kritis Mahasiswa

Kategori	Jumlah Mahasiswa	Persentase (%)
Sangat Tinggi	15	21,74
Tinggi	15	21,74
Rendah	3	4,35
Sangat Rendah	36	52,17
Total	69	100

Berdasarkan Tabel 2, kemampuan berpikir kritis mahasiswa menunjukkan distribusi yang belum merata. Sebanyak 36 mahasiswa (52,17%) berada pada kategori sangat rendah, sedangkan masing-masing 15 mahasiswa (21,74%) berada pada kategori tinggi dan sangat tinggi. Adapun kategori rendah hanya ditempati oleh 3 mahasiswa (4,35%). Distribusi ini menunjukkan bahwa lebih dari separuh responden masih memiliki kemampuan berpikir kritis yang berada di bawah kategori yang diharapkan, sementara mahasiswa yang telah mencapai kategori tinggi dan sangat tinggi masih berjumlah kurang dari setengah dari keseluruhan responden.

Dominasi kategori sangat rendah mengindikasikan bahwa sebagian besar mahasiswa masih mengalami kesulitan dalam menyelesaikan soal-soal yang menuntut keterampilan berpikir tingkat tinggi. Kondisi ini menunjukkan bahwa mahasiswa belum sepenuhnya mampu mengintegrasikan pengetahuan, menganalisis informasi secara kritis, mengevaluasi bukti, maupun memberikan justifikasi terhadap keputusan yang diambil. Kemampuan berpikir kritis pada dasarnya tidak hanya ditunjukkan melalui kemampuan mengingat atau memahami konsep, tetapi juga melalui kemampuan menggunakan pengetahuan tersebut untuk menganalisis permasalahan dan menghasilkan keputusan yang didukung alasan logis.

Penelitian (Hyytinen et al., 2025) mengenai asesmen berpikir kritis di pendidikan tinggi menunjukkan bahwa keterampilan seperti *inference* dan *evaluation*

merupakan komponen penting yang banyak digunakan dalam pengukuran berpikir kritis. Dengan demikian, rendahnya proporsi mahasiswa pada kategori tinggi perlu dipahami bukan semata-mata sebagai keterbatasan penguasaan konsep, tetapi sebagai indikasi bahwa sebagian mahasiswa belum secara konsisten mampu mengoperasikan pengetahuan dalam proses penalaran tingkat tinggi.

Dalam konteks biologi, tuntutan tersebut menjadi semakin penting karena pembelajaran biologi tidak hanya membutuhkan penguasaan fakta dan konsep, tetapi juga kemampuan menggunakan bukti untuk membangun penjelasan dan menarik kesimpulan ilmiah. Penelitian Dowd et al., (2018) menunjukkan bahwa *scientific reasoning* pada mahasiswa bidang biologi memiliki keterkaitan yang kuat dengan kemampuan *inference*. Temuan tersebut memperlihatkan bahwa kemampuan menghasilkan inferensi dari informasi atau bukti merupakan bagian penting dalam penalaran ilmiah mahasiswa biologi. Oleh karena itu, dominasi mahasiswa pada kategori sangat rendah dalam penelitian ini dapat menjadi indikasi bahwa pengalaman mahasiswa dalam menggunakan informasi dan bukti untuk menghasilkan kesimpulan atau keputusan yang beralasan masih perlu diperkuat. Interpretasi ini tetap perlu ditempatkan secara hati-hati karena penelitian ini bersifat deskriptif dan tidak menguji secara langsung faktor penyebab rendahnya kemampuan berpikir kritis.

Variasi kemampuan yang ditemukan juga dapat dikaitkan dengan karakteristik

pengalaman belajar dan asesmen yang diterima mahasiswa. Penelitian Uminski et al., (2024) terhadap 111 ujian pada 100 institusi pendidikan tinggi di Amerika Serikat menunjukkan bahwa 92% butir soal biologi berada pada tingkat *remember*, *understand*, atau *apply*, sedangkan hanya 8% yang menuntut kemampuan *analyze*, *evaluate*, atau *create*. Penelitian tersebut juga menemukan bahwa hanya 7% butir ujian yang melibatkan *scientific practices*, sementara 32% ujian tidak mengukur *scientific practices* sama sekali. Kondisi tersebut menunjukkan bahwa kesempatan mahasiswa untuk berlatih menggunakan keterampilan berpikir tingkat tinggi dalam pembelajaran dan asesmen dapat relatif terbatas apabila pengalaman belajar lebih banyak berorientasi pada penguasaan fakta dan konsep.

Di sisi lain, keberadaan 30 mahasiswa (43,48%) yang berada pada kategori tinggi dan sangat tinggi menunjukkan bahwa sebagian mahasiswa telah memiliki kemampuan berpikir kritis yang baik. Kelompok ini diperkirakan telah mampu memahami permasalahan secara komprehensif, mengevaluasi berbagai alternatif solusi, serta menyusun argumen berdasarkan bukti yang tersedia. Variasi kemampuan tersebut menunjukkan adanya heterogenitas kemampuan berpikir kritis di antara mahasiswa, sehingga pembelajaran perlu dirancang secara adaptif agar mampu mengakomodasi kebutuhan mahasiswa dengan tingkat kemampuan yang berbeda.

Heterogenitas tersebut memiliki implikasi penting bagi pembelajaran calon guru biologi. Marangio et al., (2024) menegaskan bahwa pengembangan *creative and critical thinking* pada *science pre-service teachers* memerlukan perhatian yang eksplisit dalam proses pendidikan guru dan tidak cukup hanya mengandalkan keterlibatan mahasiswa dalam pembelajaran bidang sains. Pembelajaran perlu memberikan kesempatan kepada mahasiswa untuk secara aktif menggunakan keterampilan berpikir tersebut dalam berbagai konteks. Dengan adanya variasi kemampuan yang cukup besar dalam penelitian ini, strategi pembelajaran juga

perlu dirancang secara adaptif agar mahasiswa dengan kemampuan sangat rendah memperoleh *scaffolding* yang memadai, sementara mahasiswa dengan kemampuan lebih tinggi tetap memperoleh tantangan kognitif yang sesuai.

Hasil pemetaan ini pada akhirnya memberikan gambaran bahwa pengembangan keterampilan berpikir kritis masih menjadi tantangan dalam pembelajaran biologi di perguruan tinggi. Upaya pengembangannya perlu diarahkan pada pengalaman belajar yang tidak hanya berorientasi pada penguasaan konsep, tetapi juga mendorong mahasiswa untuk menganalisis permasalahan kontekstual, mengevaluasi informasi dari berbagai sumber, menggunakan bukti untuk menarik inferensi, menyusun argumentasi ilmiah, dan merefleksikan proses pengambilan keputusan. Arah tersebut sejalan dengan temuan Lu et al., (2025) melalui *meta-analysis* yang menunjukkan bahwa pembelajaran berbasis masalah dapat mendukung pengembangan keterampilan berpikir kritis ketika mahasiswa terlibat secara aktif dalam proses pemecahan masalah. Dengan demikian, temuan penelitian ini memberikan dasar empiris bagi perlunya penguatan pembelajaran dan asesmen yang secara sistematis menyediakan pengalaman penalaran ilmiah dan berpikir tingkat tinggi bagi calon guru biologi.

Hasil pemetaan ini memberikan gambaran bahwa pengembangan kemampuan berpikir kritis masih menjadi tantangan dalam pembelajaran biologi di perguruan tinggi. Oleh karena itu, proses pembelajaran perlu lebih banyak memberikan pengalaman belajar yang mendorong mahasiswa untuk menganalisis permasalahan kontekstual, mengevaluasi informasi dari berbagai sumber, menyusun argumentasi ilmiah, serta merefleksikan proses pengambilan keputusan. Dengan demikian, mahasiswa tidak hanya menguasai konsep biologi, tetapi juga mampu menerapkan proses berpikir kritis dalam menghadapi berbagai persoalan nyata.

Disparitas Profil Kemampuan Berpikir Kritis Antar Kampus

Tabel 3. Distribusi Kemampuan Berpikir Kritis Berdasarkan Kampus

Kategori	Range	Nomor Mahasiswa	Persentase Mahasiswa Kampus A (%)	Persentase Mahasiswa Kampus B (%)
Keterampilan berpikir kritis sangat tinggi	>1.39	57B, 07A, 10A, 22A, 58B, 66B, 68B, 15A, 33A, 39A, 65B, 17A, 26A, 40A, 59B	19.5	26.08
Keterampilan berpikir kritis tinggi	0.12—1.39	34A, 51B, 56B, 64B, 60B, 47B, 25A, 37A, 38A, 45A, 53B, 42A, 43A, 44A, 35A	19.56	26.08
Keterampilan berpikir kritis rendah	-1.15—0.12	06A, 18A, 67B	4.35	4.35
Keterampilan berpikir kritis sangat rendah	< (-1.15)	23A, 28A, 29A, 32A, 48B, 50B, 02A, 61B, 16A, 30A, 49B, 62B, 08A, 19A, 20A, 24A, 54B, 12A, 14A, 31A, 01A, 04A, 05A, 13A, 27A, 03A, 11A, 52B, 63B, 69B, 21A, 41A, 55B, 09A, 36A, 46A	56.52	43.48

Catatan:

- Jumlah mahasiswa kampus A sebanyak 46 orang
- Jumlah mahasiswa kampus B sebanyak 23 orang

Distribusi *person measure* menunjukkan adanya disparitas profil kemampuan berpikir kritis antara mahasiswa Kampus A dan Kampus B. Mahasiswa Kampus B memiliki proporsi kategori tinggi dan sangat tinggi yang lebih besar dibandingkan Kampus A. Sebaliknya, proporsi mahasiswa pada kategori sangat rendah lebih banyak ditemukan pada Kampus A.

Temuan tersebut menunjukkan bahwa kemampuan berpikir kritis mahasiswa pada kedua institusi memiliki karakteristik distribusi yang berbeda. Perbedaan ini tidak dimaksudkan untuk menunjukkan bahwa salah satu kampus memiliki kemampuan berpikir kritis yang lebih baik secara mutlak, melainkan menggambarkan adanya variasi profil kemampuan mahasiswa berdasarkan hasil pemetaan menggunakan Model Rasch.

Secara konseptual, perbedaan profil tersebut sejalan dengan temuan penelitian mengenai perkembangan berpikir kritis dalam pendidikan tinggi. Penelitian Huber & Kuncel, (2016) menunjukkan bahwa keterampilan dan disposisi berpikir kritis dapat berkembang selama pengalaman pendidikan tinggi, tetapi perkembangan

tersebut tidak semata-mata ditentukan oleh keberadaan mahasiswa di perguruan tinggi. Pengalaman akademik dan karakteristik pembelajaran yang diterima mahasiswa menjadi bagian penting dalam proses perkembangan tersebut. Temuan ini memberikan dasar bahwa perbedaan profil kemampuan antar kelompok mahasiswa dapat dipahami dalam konteks variasi pengalaman pendidikan yang mereka peroleh, meskipun hubungan tersebut tidak dapat digunakan untuk menjelaskan secara kausal perbedaan Kampus A dan Kampus B dalam penelitian ini.

Berdasarkan penelusuran konteks pembelajaran di lapangan, mahasiswa Kampus B secara berkala memperoleh pengalaman belajar melalui berbagai pendekatan inovatif, seperti *Problem-Based Learning* (PBL), *Project-Based Learning* (PjBL), *Discovery Learning*, integrasi Ekopedagogi-PjBL, *Education for Sustainable Development* (ESD) berbasis *Deep Learning*, serta *Reading, Questioning, and Answering* (RQA). Sebaliknya, proses pembelajaran di Kampus A lebih banyak didominasi oleh presentasi menggunakan media *PowerPoint* dan diskusi kelas. Variasi

pengalaman belajar tersebut berpotensi memberikan kesempatan yang berbeda bagi mahasiswa dalam mengembangkan kemampuan analisis, evaluasi, pemecahan masalah, dan refleksi. Namun demikian, penelitian ini tidak menguji secara langsung pengaruh model pembelajaran terhadap kemampuan berpikir kritis sehingga penjelasan tersebut hanya digunakan sebagai interpretasi kontekstual terhadap temuan yang diperoleh.

Meskipun Kampus B menunjukkan performa yang lebih baik dibandingkan Kampus A, persentase mahasiswa pada kategori "Sangat Rendah" masih tergolong tinggi, yaitu 43,48%. Temuan ini mengindikasikan adanya polarisasi kemampuan berpikir kritis di dalam kelas. Dari perspektif *Cognitive Load Theory*, penerapan berbagai model pembelajaran aktif secara simultan dapat meningkatkan keterlibatan kognitif, tetapi juga berpotensi menimbulkan *cognitive overload* pada mahasiswa yang memiliki pengetahuan awal, regulasi diri, dan kemandirian belajar yang masih terbatas. Akibatnya, sebagian mahasiswa mengalami kesulitan mengonstruksi penalaran kritis secara mendalam meskipun telah mengikuti berbagai aktivitas pembelajaran. Selain itu, ketergantungan pada LKM (lembar kerja mahasiswa) dan *scaffolding* selama perkuliahan diduga menyebabkan sebagian mahasiswa belum mampu mentransfer strategi berpikir secara mandiri ketika

menghadapi tes individual. Temuan ini menunjukkan bahwa keberhasilan pembelajaran aktif tidak hanya ditentukan oleh variasi model pembelajaran, tetapi juga oleh kesiapan kognitif mahasiswa serta pemberian scaffolding yang terukur dan bertahap (Gkintoni et al., 2025; Michalsky, 2024; Vendrell & Johnston, 2026).

Karakteristik Tingkat Kesulitan Butir Berdasarkan Indikator Berpikir Kritis

Selain memetakan kemampuan berpikir kritis mahasiswa, model *Rasch* juga menghasilkan informasi mengenai tingkat kesulitan setiap butir soal (*item measure*). Tingkat kesulitan butir menunjukkan seberapa besar kemampuan yang diperlukan mahasiswa untuk menjawab suatu soal dengan benar. Semakin tinggi nilai logit suatu butir, semakin sulit butir tersebut dijawab oleh responden, sedangkan nilai logit yang semakin rendah menunjukkan bahwa butir tersebut relatif lebih mudah dijawab.

Hasil analisis menunjukkan bahwa nilai rata-rata (*mean*) kesulitan butir adalah 0,00 logit dengan standar deviasi sebesar 0,90 logit. Berdasarkan nilai tersebut, butir soal dikelompokkan menjadi empat kategori, yaitu sangat sulit, sulit, mudah, dan sangat mudah. Kriteria pengelompokan tingkat kesulitan butir disajikan pada Tabel 4.

Tabel 4. Distribusi Tingkat Kesulitan Butir

Kategori	Range	Nomor Soal
Soal kritis dengan kategori sangat sulit	>0.90	S25, S3, S26
Soal kritis dengan kategori sulit	0.00—0.90	S10, S12, S8, S13, S23, S20, S22, S9, S28, S7
Soal kritis dengan kategori mudah	-0.90—0.00	S15, S2, S6, S17, S21, S16, S29, S30, S27, S19, S24, S18
Soal kritis dengan kategori sangat mudah	< (-0.90)	S4, S14, S5, S11, S1

Berdasarkan Tabel 4, terdapat 3 butir (10,00%) yang termasuk kategori sangat sulit, 10 butir (33,33%) berkategori sulit, 12 butir (40,00%) berkategori mudah, dan 5 butir (16,67%) berkategori sangat mudah. Sebaran tersebut menunjukkan bahwa instrumen memiliki variasi tingkat kesulitan yang baik sehingga mampu mengukur kemampuan berpikir kritis mahasiswa pada

berbagai tingkat kemampuan. Keberadaan butir dengan tingkat kesulitan yang beragam juga menunjukkan bahwa instrumen tidak hanya efektif dalam mengidentifikasi mahasiswa dengan kemampuan tinggi maupun rendah, tetapi juga mampu membedakan tingkat kemampuan secara lebih akurat.

Untuk memperoleh gambaran yang lebih spesifik mengenai aspek berpikir kritis yang diukur, tingkat kesulitan butir selanjutnya dianalisis berdasarkan kerangka

indikator berpikir kritis milik Facione. Hasil pengelompokan tersebut disajikan pada Tabel 5.

Tabel 5. Karakteristik Tingkat Kesulitan Butir Berdasarkan Indikator Keterampilan Berpikir Kritis

Indikator	Nomor Butir	Butir Sangat Sulit	Butir Sulit	Butir Mudah	Butir Sangat mudah
Interpretasi	7, 14, 16, 22, 23, 30	-	7, 22, 23	16, 30	14
Analisis	2, 6, 10, 15, 18, 19, 28, 29	-	10, 28	2, 6, 15, 18, 19, 29	-
Inferensi	3, 9, 20	3	9, 20	-	-
Evaluasi	1, 5, 8, 11, 12, 17, 21, 24, 25, 26, 27	25, 26	8, 12	17, 21, 24, 27	1, 5, 11
Eksplanasi	-	-	-	-	-
Pengaturan Diri	4, 13	-	13	-	4

Berdasarkan Tabel 4, dari 30 butir keterampilan berpikir kritis terdapat 3 butir sangat sulit, 10 butir sulit, 12 butir mudah, dan 5 butir sangat mudah. Ketika dipetakan berdasarkan indikator Facione pada Tabel 5, tampak adanya perbedaan distribusi tingkat kesulitan antarindikator. Perbedaan tersebut menunjukkan bahwa tingkat kesulitan butir tidak hanya berkaitan dengan posisi *item measure*, tetapi juga dengan tuntutan kognitif yang melekat pada tugas yang diberikan. Hal ini sejalan dengan kajian Hyytinen et al., (2025) yang menegaskan bahwa asesmen *critical thinking* melibatkan keterampilan kognitif yang beragam, dan evaluation serta inference merupakan aspek yang paling sering menjadi sasaran pengukuran.

Indikator interpretasi terdiri atas enam butir (S7, S14, S16, S22, S23, dan S30), dengan tiga butir sulit (S7, S22, S23), dua butir mudah (S16, S30), dan satu butir sangat mudah (S14). Pola ini menunjukkan bahwa kemampuan interpretasi memiliki tingkat kesulitan yang bervariasi. Butir interpretasi yang sulit umumnya menuntut mahasiswa memahami makna informasi dalam konteks yang lebih kompleks, sedangkan butir yang sangat mudah lebih dekat pada pengenalan atau pengelompokan informasi secara langsung. Dengan demikian, interpretasi tidak dapat dipandang sebagai kemampuan yang selalu rendah atau tinggi kesulitannya, karena kompleksitas konteks dan informasi pada masing-masing butir turut menentukan beban kognitifnya.

Indikator analisis mencakup delapan butir (S2, S6, S10, S15, S18, S19, S28, dan

S29), dan enam di antaranya berada pada kategori mudah, sedangkan dua butir (S10 dan S28) berada pada kategori sulit. Dominasi butir mudah menunjukkan bahwa sebagian besar aktivitas analitis dalam instrumen relatif dapat diselesaikan mahasiswa. Namun, S10 dan S28 menjadi pengecualian karena menuntut mahasiswa mengidentifikasi kelemahan atau kesalahan dalam struktur argumen, sehingga tuntutan penalarannya lebih tinggi. Temuan ini konsisten dengan Comstock & Grün, (2026) yang menunjukkan bahwa asesmen pilihan ganda dapat mengukur *critical reasoning* secara lebih mendalam ketika butir menuntut identifikasi hubungan antara alasan, kesimpulan, keberatan, dan kualitas argumen.

Indikator inferensi merupakan pola yang paling menonjol karena ketiga butirnya (S3, S9, dan S20) berada pada kategori sangat sulit atau sulit. Tidak terdapat butir inferensi pada kategori mudah maupun sangat mudah. Pola ini menunjukkan bahwa kemampuan menarik kesimpulan berdasarkan informasi yang tersedia merupakan tuntutan kognitif yang relatif tinggi pada instrumen ini. Temuan tersebut sejalan dengan Paulsen & Kolstø, (2022) yang menemukan bahwa siswa dapat mengalami kesulitan ketika harus menghubungkan bukti atau premis dengan kesimpulan secara tepat; bahkan 28% respons yang mereka analisis menunjukkan penalaran induktif yang tampak kuat tetapi menghasilkan jawaban salah karena didasarkan pada premis yang tidak relevan atau tidak sesuai konteks. Dengan demikian,

tingginya *item measure* pada butir inferensi dalam penelitian ini dapat dipahami sebagai indikasi bahwa penarikan kesimpulan memerlukan integrasi informasi dan penggunaan bukti secara lebih hati-hati.

Evaluasi merupakan indikator dengan jumlah butir terbanyak, yaitu 11 butir (S1, S5, S8, S11, S12, S17, S21, S24, S25, S26, dan S27), sekaligus memiliki rentang tingkat kesulitan paling luas. Terdapat dua butir sangat sulit (S25 dan S26), dua butir sulit (S8 dan S12), empat butir mudah (S17, S21, S24, dan S27), serta tiga butir sangat mudah (S1, S5, dan S11). Sebaran tersebut menunjukkan bahwa keterampilan evaluasi dapat diukur melalui tugas dengan tingkat kompleksitas yang sangat beragam. Variasi ini sesuai dengan temuan Gu et al., (2021) yang menunjukkan bahwa karakteristik *item critical thinking* dapat dipengaruhi oleh keterampilan kognitif yang ditargetkan, tipe respons, tuntutan membaca, dan karakteristik penyajian asesmen. Karena itu, perbedaan tingkat kesulitan pada butir evaluasi tidak secara otomatis berarti kemampuan evaluasi mahasiswa lebih rendah, tetapi juga dapat mencerminkan perbedaan kompleksitas tugas pada masing-masing butir.

Indikator eksplanasi tidak terwakili dalam 30 butir. Kondisi ini perlu dicatat sebagai karakteristik instrumen, bukan sebagai indikasi bahwa mahasiswa tidak memiliki kemampuan eksplanasi. Sebagian besar soal merupakan *selected-response items* yang meminta mahasiswa memilih alternatif jawaban, sehingga kesempatan untuk mengungkapkan hasil penalaran, memberikan justifikasi, atau menjelaskan dasar suatu keputusan sangat terbatas.

Pengaturan diri hanya terdiri atas dua butir, yaitu S4 berkategori sangat mudah dan S13 berkategori sulit. Sebaran yang kontras ini menunjukkan bahwa keterampilan regulasi diri belum cukup terwakili untuk menggambarkan pola kesulitan secara konsisten. Selain itu, karena jumlah butirnya sangat terbatas, interpretasi terhadap indikator ini perlu dilakukan secara hati-hati. Temuan tersebut juga memperlihatkan pentingnya kesesuaian antara tujuan pengukuran, konstruk yang hendak diwakili, dan desain item, sebagaimana ditekankan dalam kajian asesmen *critical thinking* di pendidikan tinggi.

Jika keenam indikator tersebut dilihat secara bersamaan, inferensi menjadi indikator dengan kecenderungan kesulitan tertinggi, karena seluruh butirnya berada pada kategori sulit dan sangat sulit. Sebaliknya, analisis didominasi butir mudah, sedangkan evaluasi menunjukkan rentang kesulitan paling luas. Temuan ini mendukung penggunaan *item measure Rasch* sebagai informasi diagnostik karena posisi relatif setiap butir dapat digunakan untuk melihat variasi tuntutan kemampuan yang diperlukan untuk menjawabnya.

Lebih lanjut, pola tersebut menunjukkan bahwa butir sulit tidak selalu identik dengan satu indikator tertentu. Tingkat kesulitan juga dipengaruhi oleh kompleksitas stimulus, jenis penalaran, tuntutan membaca, struktur argumen, serta jenis respons yang diperlukan. Hal tersebut sejalan dengan penelitian Gu et al., (2021) yang menemukan bahwa karakteristik *item* dan mode asesmen dapat memengaruhi kesulitan dan fungsi item. Oleh karena itu, hasil Tabel 5 lebih tepat digunakan untuk mengidentifikasi kecenderungan tuntutan kognitif setiap indikator, bukan untuk menyatakan bahwa suatu indikator secara absolut lebih sulit daripada indikator lainnya.

Secara pedagogis, hasil tersebut menunjukkan bahwa pembelajaran perlu memberikan perhatian lebih pada inferensi, terutama latihan menghubungkan bukti dengan kesimpulan, mempertimbangkan alternatif, dan menghindari kesimpulan yang tidak didukung informasi. Pada indikator analisis, dominasi butir mudah dapat menjadi dasar untuk mengembangkan tugas analitis yang lebih kompleks. Sementara itu, variasi luas pada evaluasi menunjukkan perlunya menyediakan tugas evaluatif dengan tingkat kompleksitas bertahap. Pendekatan ini sejalan dengan pandangan bahwa asesmen berpikir kritis perlu dirancang berdasarkan keterampilan spesifik yang hendak diukur dan karakteristik tugas yang digunakan.

Differential Item Functioning (DIF)

Selain memetakan profil kemampuan berpikir kritis mahasiswa, analisis DIF dilakukan untuk mengidentifikasi apakah terdapat butir yang memiliki karakteristik berbeda ketika dikerjakan oleh mahasiswa dari Kampus A dan Kampus B. Analisis ini

bertujuan untuk mengetahui apakah suatu butir memberikan tingkat kesulitan yang relatif berbeda pada kedua kelompok meskipun mengukur konstruk kemampuan

berpikir kritis yang sama. Suatu butir dinyatakan mengalami DIF apabila memiliki nilai probabilitas (*p-value*) kurang dari 0,05.

Tabel 6. Butir yang Mengalami *Differential Item Functioning* (DIF)

Nomor Soal	Nilai Kemampuan A (Logit)	Nilai Kemampuan B (Logit)	DIF Contrast	Probabilitas (p)	Kelompok yang Diuntungkan
S19	-0,85	0,80	-1,64	0,0150	Kampus A (Lebih Mudah)
S27	-0,85	0,80	-1,64	0,0150	Kampus A (Lebih Mudah)

Berdasarkan Tabel 6, terdapat dua butir yang teridentifikasi mengalami *Differential Item Functioning* (DIF), yaitu S19 dan S27. Kedua butir memiliki nilai *p* sebesar 0,015 ($<0,05$), sehingga menunjukkan adanya perbedaan tingkat kesulitan butir antara mahasiswa Kampus A dan Kampus B. Nilai *DIF contrast* yang negatif mengindikasikan bahwa kedua butir tersebut relatif lebih mudah dijawab oleh mahasiswa Kampus A dibandingkan mahasiswa Kampus B.

Temuan ini menarik karena berbeda dengan hasil pemetaan kemampuan secara keseluruhan yang menunjukkan bahwa distribusi kemampuan berpikir kritis mahasiswa Kampus B cenderung lebih tinggi dibandingkan Kampus A. Dengan demikian, perbedaan kemampuan berpikir kritis antar kampus tidak bersifat seragam pada seluruh indikator. Mahasiswa Kampus B memiliki profil kemampuan yang lebih baik secara umum, namun pada butir tertentu mahasiswa Kampus A justru menunjukkan performa yang lebih baik.

Kedua butir yang mengalami DIF berada pada indikator evaluasi. Hal ini menunjukkan bahwa perbedaan antar kampus terutama tampak pada kemampuan mengevaluasi informasi, menilai kualitas argumen, serta memberikan pertimbangan terhadap alternatif solusi berdasarkan bukti yang tersedia. Meskipun mahasiswa Kampus B lebih sering memperoleh pengalaman belajar melalui berbagai pendekatan inovatif, seperti *Problem-Based Learning* (PBL), *Project-Based Learning* (PjBL), *Discovery Learning*, integrasi Ekopedagogi-PjBL, *Education for Sustainable Development* (ESD), *Deep Learning*, dan *Reading*,

Questioning, and Answering (RQA), hasil ini mengindikasikan bahwa pengalaman belajar tersebut belum tentu menghasilkan keunggulan yang merata pada seluruh indikator berpikir kritis.

Sebaliknya, mahasiswa Kampus A yang lebih banyak mengikuti pembelajaran berbasis presentasi dan diskusi kelas justru menunjukkan keunggulan pada dua butir tersebut. Temuan ini mengindikasikan bahwa kemampuan mengevaluasi informasi kemungkinan dipengaruhi oleh berbagai faktor lain, seperti karakteristik pembelajaran pada mata kuliah tertentu, pengalaman akademik mahasiswa, maupun konteks penyajian soal. Oleh karena itu, hasil DIF tidak dapat diinterpretasikan sebagai bukti bahwa salah satu model pembelajaran lebih efektif dibandingkan model lainnya, melainkan menunjukkan adanya variasi penguasaan pada indikator tertentu yang memerlukan kajian lebih lanjut.

Perbedaan fungsi pada S19 dan S27 juga perlu ditafsirkan secara hati-hati karena DIF tidak secara otomatis berarti bahwa suatu butir merupakan butir yang buruk. DIF menunjukkan bahwa karakteristik respons terhadap butir tersebut berbeda antarkelompok setelah mempertimbangkan kemampuan pada konstruk yang diukur. Penelitian mengenai pengembangan dan evaluasi instrumen juga menunjukkan bahwa identifikasi DIF perlu diikuti dengan pemeriksaan terhadap isi, konteks, bahasa, maupun karakteristik kognitif butir untuk menemukan kemungkinan sumber perbedaannya. Dalam penelitian Kurnaz & Yildiz, (2023), misalnya, sebagian sumber DIF ditemukan berkaitan dengan isi butir dan sebagian lainnya berkaitan dengan bahasa

atau ekspresi. Dengan demikian, S19 dan S27 sebaiknya tidak langsung dieliminasi hanya berdasarkan signifikansi statistik, tetapi ditelaah kembali berdasarkan konteks soal dan tuntutan kognitif yang terkandung di dalamnya.

Secara keseluruhan, hasil analisis DIF melengkapi temuan mengenai disparitas kemampuan berpikir kritis mahasiswa. Analisis ini menunjukkan bahwa perbedaan profil antar kampus tidak hanya terlihat pada distribusi kemampuan mahasiswa, tetapi juga pada karakteristik respons terhadap butir tertentu. Informasi tersebut dapat dimanfaatkan sebagai bahan evaluasi dalam pengembangan pembelajaran maupun penyempurnaan instrumen agar mampu mengukur kemampuan berpikir kritis secara lebih adil pada berbagai kelompok mahasiswa.

Implikasi Temuan Penelitian

Hasil penelitian menunjukkan bahwa analisis Rasch mampu memberikan gambaran yang komprehensif mengenai profil kemampuan berpikir kritis mahasiswa. Selain memetakan distribusi kemampuan secara keseluruhan, analisis ini juga mengidentifikasi adanya disparitas profil antar kampus, variasi tingkat kesulitan butir pada setiap indikator berpikir kritis, serta karakteristik respons mahasiswa terhadap instrumen yang digunakan. Informasi tersebut dapat dimanfaatkan sebagai dasar dalam merancang strategi pembelajaran yang lebih adaptif sesuai dengan kebutuhan mahasiswa.

Temuan ini juga menunjukkan bahwa peningkatan kemampuan berpikir kritis tidak cukup dilakukan melalui penyampaian materi, tetapi memerlukan pengalaman belajar yang memberikan kesempatan kepada mahasiswa untuk menganalisis permasalahan autentik, mengevaluasi informasi secara kritis, menyusun argumentasi berbasis bukti, dan merefleksikan proses pengambilan keputusan.

Penelitian Liu & Pásztor, (2022) melalui meta-analysis terhadap 50 penelitian empiris pada pendidikan tinggi, menunjukkan bahwa *Problem-Based Learning* (PBL) memberikan pengaruh yang kuat terhadap *overall critical thinking*, *critical-thinking skills*, dan *critical-thinking*

disposition. Temuan tersebut memperkuat implikasi penelitian ini bahwa pembelajaran perlu menyediakan aktivitas yang menempatkan mahasiswa sebagai pelaku aktif dalam proses penalaran dan pemecahan masalah, bukan hanya sebagai penerima informasi.

Dalam konteks pembelajaran biologi, implikasi tersebut dapat diwujudkan melalui penggunaan permasalahan autentik yang berkaitan dengan fenomena biologis dan kehidupan sehari-hari. Mahasiswa dapat diarahkan untuk mengidentifikasi masalah, menginterpretasikan data, membandingkan berbagai bukti, mengevaluasi validitas suatu penjelasan, serta menyusun kesimpulan berdasarkan bukti ilmiah. Pendekatan tersebut menjadi semakin relevan karena pembelajaran sains menuntut mahasiswa tidak hanya memahami konsep, tetapi juga menggunakan konsep tersebut untuk menjelaskan fenomena dan memecahkan permasalahan. Oleh karena itu, hasil pemetaan ini dapat menjadi dasar bagi dosen dalam merancang pembelajaran dan asesmen yang lebih efektif untuk mengembangkan kemampuan berpikir kritis mahasiswa.

Selain strategi pembelajaran, temuan penelitian juga memiliki implikasi terhadap asesmen. Variasi tingkat kesulitan butir yang ditemukan menunjukkan pentingnya menyediakan tugas dan soal yang memiliki tuntutan kognitif beragam. Asesmen sebaiknya tidak didominasi oleh pertanyaan yang hanya menuntut mahasiswa mengingat atau memahami informasi, tetapi juga mencakup tugas yang meminta mahasiswa menganalisis informasi, mengevaluasi bukti, memberikan justifikasi, dan menghasilkan keputusan. Dalam konteks mata kuliah fisiologi, Anderson & Fernandez-Branson, (2024) menunjukkan pentingnya penyesuaian antara tujuan pembelajaran, asesmen, dan *active learning* dalam upaya mempromosikan *critical thinking*. Temuan tersebut mendukung perlunya keselarasan antara apa yang ingin dikembangkan melalui pembelajaran dengan apa yang benar-benar diukur melalui asesmen.

Simpulan

Berdasarkan analisis menggunakan Rasch Model, instrumen tes berpikir kritis yang digunakan memiliki kualitas,

reliabilitas, dan konsistensi yang baik sehingga layak digunakan untuk mengukur keterampilan berpikir kritis mahasiswa calon guru biologi. Hasil penelitian menunjukkan adanya disparitas kemampuan berpikir kritis antara kedua kampus. Sebagian besar mahasiswa Kampus A berada pada kategori sangat rendah, sedangkan lebih dari separuh mahasiswa Kampus B berada pada kategori tinggi dan sangat tinggi. Temuan ini mengindikasikan bahwa lingkungan belajar yang mendukung pembelajaran aktif dan berpusat pada mahasiswa berpotensi memberikan kontribusi positif terhadap pengembangan keterampilan berpikir kritis.

Hambatan utama mahasiswa pada kedua kampus terletak pada kemampuan berpikir kritis tingkat tinggi, terutama inferensi, evaluasi argumen ilmiah, dan pengambilan keputusan berbasis bukti. Analisis *Differential Item Functioning* (DIF) juga menunjukkan bahwa mahasiswa Kampus A memiliki keunggulan dalam mengevaluasi informasi, menilai kualitas argumen, serta memberikan pertimbangan terhadap alternatif solusi berdasarkan bukti yang tersedia. dibandingkan mahasiswa Kampus B.

Hasil penelitian ini mengimplikasikan perlunya penguatan pembelajaran yang lebih berorientasi pada pemecahan masalah, analisis data, evaluasi argumen ilmiah, dan isu-isu sosiosaintifik. Selain itu, institusi pendidikan perlu menciptakan lingkungan belajar yang lebih interaktif dan mendukung pengembangan keterampilan berpikir tingkat tinggi sebagai bekal calon guru biologi menghadapi tantangan abad ke-21.

Ucapan Terima Kasih (usahakan ada)

Kami ingin mengucapkan terima kasih kepada Beasiswa Pendidikan Indonesia (BPI), Badan Pembiayaan Pendidikan Tinggi (BPPT), Pusat Pembiayaan dan Penilaian Pendidikan Tinggi (PPAPT) di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, serta Lembaga Pengelola Dana Pendidikan (LPDP) atas pendanaan dan dukungan yang diberikan untuk penelitian dan penerbitan artikel ini.

Daftar Pustaka

Aflalo, E. (2026). Student Involvement in Digital Tool Selection: A Pedagogical

Approach to Critical Thinking-Oriented Learning. *Education Sciences*, 16(4), 512.

<https://doi.org/10.3390/educsci16040512>

Anderson, L. C., & Fernandez-Branson, C. (2024). Alignment of learning objectives, assessments, and active learning to promote critical thinking in a first-year medical physiology course: lessons learned. *Advances in Physiology Education*, 48(2), 385–394. <https://doi.org/10.1152/advan.00096.2023>

Anggrella, D. P., & Sudrajat, A. K. (2025). Profile and Views on Botanical Literacy: A Study From a Final Year Preservice Biology Teacher. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, 11(2), 548–561. <https://doi.org/10.22219/jpbi.v11i2.40470>

Apriliani, E. A., Afandi, Yuniarti, A., & Bilanti, E. (2022). Critical Thinking Appraisal Scale Profile of Prospective Biology Teachers in Tanjungpura University. *The 3rd International Conference on Science Education*, 070022. <https://doi.org/10.1063/5.0113151>

Bond, T., Yan, Z., & Heene, M. (2020). *Applying the Rasch Model*. Routledge. <https://doi.org/10.4324/9780429030499>

Comstock, G., & Grün, D. (2026). Assessing critical thinking about argumentative essays: the Critical Reasoning and Inference Test (CRIT). *Assessment & Evaluation in Higher Education*, 1–14. <https://doi.org/10.1080/02602938.2026.2653881>

Darussyamsu, R., Lufri, L., Ahda, Y., Alberida, H., Ambiyar, A., & Samsudin, S. (2024). The Effectiveness of Project-Based Learning Model with RAHMA Syntax to Improve Prospective Biology Teachers' Critical Thinking Skills. *Jurnal Pendidikan IPA Indonesia*, 13(4). <https://doi.org/10.15294/xy8mr440>

Dowd, J. E., Thompson, R. J., Schiff, L. A., & Reynolds, J. A. (2018). Understanding the Complex

- Relationship between Critical Thinking and Science Reasoning among Undergraduate Thesis Writers. *CBE—Life Sciences Education*, 17(1), ar4. <https://doi.org/10.1187/cbe.17-03-0052>
- García-Carmona, A. (2025). Scientific Thinking and Critical Thinking in Science Education. *Science & Education*, 34(1), 227–245. <https://doi.org/10.1007/s11191-023-00460-5>
- Gkintoni, E., Antonopoulou, H., Sortwell, A., & Halkiopoulos, C. (2025). Challenging Cognitive Load Theory: The Role of Educational Neuroscience and Artificial Intelligence in Redefining Learning Efficacy. *Brain Sciences*, 15(2), 203. <https://doi.org/10.3390/brainsci15020203>
- Gu, L., Ling, G., Liu, O. L., Yang, Z., Li, G., Kardanova, E., & Loyalka, P. (2021). Examining mode effects for an adapted Chinese critical thinking assessment. *Assessment & Evaluation in Higher Education*, 46(6), 879–893. <https://doi.org/10.1080/02602938.2020.1836121>
- Huber, C. R., & Kuncel, N. R. (2016). Does College Teach Critical Thinking? A Meta-Analysis. *Review of Educational Research*, 86(2), 431–468. <https://doi.org/10.3102/0034654315605917>
- Hyttinen, H., Jämsä, M., Tuononen, T., & Kleemola, K. (2025). A systematic-narrative review of performance-based assessments of critical thinking in higher education. *Assessment & Evaluation in Higher Education*, 50(8), 1293–1310. <https://doi.org/10.1080/02602938.2025.2553341>
- Kurnaz, F. B., & Yildiz, H. (2023). Investigating the sources of differential item functioning: A sample critical thinking motivation scale. *International Journal of Assessment Tools in Education*, 10(3), 434–453. <https://doi.org/10.21449/ijate.1279152>
- Liu, Y., & Pásztor, A. (2022). Effects of problem-based learning instructional intervention on critical thinking in higher education: A meta-analysis. *Thinking Skills and Creativity*, 45, 101069. <https://doi.org/10.1016/j.tsc.2022.101069>
- Lu, L., Mustakim, S. S., & Muhamad, M. M. (2025). A Meta-analysis of the Effectiveness of Problem-based Learning on Critical Thinking. *European Journal of Educational Research*, 14(3), 789–804. <https://doi.org/10.12973/eu-jer.14.3.789>
- Madieva, A., Chaldanbaeva, A., Dzhaanbaeva, A., & Ckukun kyzy, N. (2025). Improving The Methodology for Training Future Biology Teachers in A Distance Learning Environment. *Social Sciences & Humanities Open*, 12, 102283. <https://doi.org/10.1016/j.ssaho.2025.102283>
- Marangio, K., Carpendale, J., Cooper, R., & Mansfield, J. (2024). Supporting the Development of Science Pre-service Teachers' Creativity and Critical Thinking in Secondary Science Initial Teacher Education. *Research in Science Education*, 54(1), 65–81. <https://doi.org/10.1007/s11165-023-10104-x>
- Maryuningsih, Y., Hidayat, T., Riandi, R., & Rustaman, N. Y. (2020). The Critical Thinking Skills of Biology Teacher Candidates Toward The Ethical Issues. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, 6(1), 65–74. <https://doi.org/10.22219/jpbi.v6i1.10779>
- Michalsky, T. (2024). Metacognitive Scaffolding for Preservice Teachers' Self-Regulated Design of Higher Order Thinking Tasks. *Heliyon*, 10(2), e24280. <https://doi.org/10.1016/j.heliyon.2024.e24280>
- Paulsen, V. H., & Kolstø, S. D. (2022). Students' reasoning when faced with test items of challenging aspects of critical thinking. *Thinking Skills and Creativity*, 43, 100969. <https://doi.org/10.1016/j.tsc.2021.100969>
- Rusch, T., Lowry, P. B., Mair, P., & Treiblmaier, H. (2017). Breaking free

- from the limitations of classical test theory: Developing and measuring information systems scales using item response theory. *Information & Management*, 54(2), 189–203. <https://doi.org/10.1016/j.im.2016.06.005>
- Sari, T. N. I., & Rakhmawati, A. (2024). Analysis The Quality of Critical Thinking and Creativity Questions in High School Biology Subjects with The Rasch Model. *Research and Development in Education (RaDEn)*, 4(1), 602–616. <https://doi.org/10.22219/raden.v4i1.32758>
- Suwita, S., Saputro, S., Sajidan, S., & Sutarno, S. (2024). Assessing Lower-Secondary School Students' Critical Thinking Skills in Photosynthesis: A Rasch Model Approach. *Journal of Baltic Science Education*, 23(6), 1278–1290. <https://doi.org/10.33225/jbse/24.23.1278>
- Syahfitri, J. (2023). Analysis of The Critical Thinking Style of Prospective Biology Teachers Students. *Jurnal Penelitian Pendidikan IPA*, 9(10), 8146–8151. <https://doi.org/10.29303/jppipa.v9i10.3704>
- Uminski, C., Burbach, S. M., & Couch, B. A. (2024). Undergraduate Biology Lecture Courses Predominantly Test Facts about Science Rather than Scientific Practices. *CBE—Life Sciences Education*, 23(2). <https://doi.org/10.1187/cbe.23-12-0244>
- Vendrell, M., & Johnston, S.-K. (2026). Scaffolding Critical Thinking With Generative AI: Design Principles for Integrating Large Language Models in Higher Education. *Computers and Education: Artificial Intelligence*, 10, 100572. <https://doi.org/10.1016/j.caeai.2026.100572>
- Yuan, R., & Liao, W. (2023). Critical Thinking in Teacher Education: Where Do We Stand And Where Can We Go? *Teachers and Teaching*, 29(6), 543–552. <https://doi.org/10.1080/13540602.2023.2252688>