Jurnal Pustaka Cendekia Pendidikan
Volume 04, No. 01, Mei– Agustus 2026, pp. 2259-2275
E-ISSN: 2988-7720
Website: https://pcpendidikan.org/index.php/jpcp

Analisis Psikometrik Modified AI Literacy Scale (MAILS) Menggunakan Model Rasch pada Mahasiswa di Sulawesi Tenggara

Minarti Usman

Universitas Halu Oleo, Kendari Indonesia
Minarti.usman@uho.ac.id

Abstract

This study aimed to analyze the psychometric quality of the Modified AI Literacy Scale (MAILS) using the Rasch Model. A quantitative approach with a psychometric analysis design was employed to evaluate 34 instrument items adapted from the MAILS, which were administered to 102 university students. Data were analyzed using Winsteps software by examining instrument reliability, separation indices, item fit, item difficulty, person fit, and the Wright Map. The results indicated that the instrument demonstrated excellent psychometric properties, with both person reliability and item reliability of 0.95, and a Cronbach's alpha ranging from 0.96 to 0.97. The person separation and item separation indices showed that the instrument effectively distinguished respondents' ability levels and item difficulty levels. A total of 25 items (73.53%) met the fit criteria, 5 items (14.71%) were classified as borderline fit, and 4 items (11.76%) were identified as misfit, indicating the need for further refinement. Item difficulty ranged from −1.02 to 1.64 logits, suggesting that the instrument is capable of measuring AI literacy across varying levels of student ability. Overall, the Modified AI Literacy Scale (MAILS) demonstrated satisfactory validity and reliability and is suitable for assessing AI literacy among university students in Indonesian higher education.

Keywords: Artificial Intelligence Literacy, MAILS, Rasch Model, Validity, Reliability, Psychometrics.

Abstrak

Penelitian ini bertujuan menganalisis kualitas instrumen Modified AI Literacy Scale (MAILS) menggunakan Model Rasch. Penelitian menggunakan pendekatan kuantitatif dengan desain analisis psikometrik terhadap 34 butir instrumen yang diadaptasi dari MAILS dan diujikan kepada 102 mahasiswa. Analisis dilakukan menggunakan perangkat lunak Winsteps dengan mengevaluasi reliabilitas instrumen, separation index, item fit, tingkat kesulitan item, person fit, dan Wright Map. Hasil penelitian menunjukkan bahwa instrumen memiliki kualitas psikometrik yang sangat baik dengan person reliability dan item reliability masing-masing sebesar 0,95, serta Cronbach Alpha sebesar 0,96–0,97. Nilai person separation dan item separation menunjukkan bahwa instrumen mampu membedakan kemampuan responden dan tingkat kesulitan item secara baik. Sebanyak 25 item (73,53%) memenuhi kategori fit, 5 item (14,71%) berada pada kategori borderline, dan 4 item (11,76%) tergolong misfit sehingga memerlukan penyempurnaan. Tingkat kesulitan item berada pada rentang −1,02 hingga 1,64 logit, yang menunjukkan bahwa instrumen mampu mengukur literasi AI pada berbagai tingkat kemampuan mahasiswa. Secara keseluruhan, Modified AI Literacy Scale (MAILS) memiliki validitas dan reliabilitas yang baik serta layak digunakan sebagai instrumen untuk mengukur literasi AI mahasiswa di pendidikan tinggi Indonesia.

Kata kunci: Artificial Intelligence Literacy, MAILS, Model Rasch, Validitas, Reliabilitas, Psikometri

✉ Corresponding author: Minarti Usman

Email Address: Minarti.usman@uho.ac.id (Universitas Halu Oleo, Kendari Indonesia)
Received 20 Juli 2026, Accepted 29 Jul 2026, Published 30 Juli 2026

PENDAHULUAN

Perkembangan teknologi berbasis kecerdasan buatan atau *Artificial Intelligence* (AI) telah membawa perubahan besar dalam berbagai sektor kehidupan, termasuk pendidikan tinggi. Kehadiran teknologi AI seperti ChatGPT, Google Gemini, dan Microsoft Copilot memungkinkan mahasiswa memperoleh akses informasi secara cepat, menghasilkan teks otomatis, membantu analisis data, hingga mendukung penyelesaian tugas akademik secara lebih efisien. Perkembangan tersebut menunjukkan

bahwa AI telah menjadi bagian penting dalam proses pembelajaran dan kehidupan akademik mahasiswa di era digital (Kasneci et al., 2023). (Unesco, 2024) menegaskan bahwa AI berpotensi meningkatkan kualitas pendidikan apabila digunakan secara etis, bertanggung jawab, dan berpusat pada manusia.

Di lingkungan perguruan tinggi, penggunaan AI memberikan peluang besar dalam meningkatkan kualitas pembelajaran, kreativitas, dan produktivitas mahasiswa. Namun demikian, penggunaan AI juga menimbulkan berbagai tantangan, terutama berkaitan dengan kemampuan mahasiswa dalam memahami cara kerja AI, mengevaluasi keakuratan informasi yang dihasilkan, menggunakan AI secara etis, serta menghindari ketergantungan berlebihan terhadap teknologi otomatis. Penelitian yang dilakukan oleh Astrid Carolus dan rekan-rekannya menunjukkan bahwa literasi AI tidak hanya berkaitan dengan kemampuan menggunakan AI, tetapi juga mencakup kemampuan memahami AI, mendeteksi penggunaan AI, serta memahami etika AI dalam kehidupan sehari-hari dan Pendidikan (Carolus et al., 2023). Oleh karena itu, literasi AI (*AI literacy*) menjadi kompetensi penting yang perlu dimiliki mahasiswa agar mampu menggunakan teknologi AI secara bijak, kritis, dan bertanggung jawab.

Konsep literasi AI berkembang sebagai respons terhadap meningkatnya penggunaan teknologi AI dalam kehidupan sehari-hari dan dunia pendidikan. Menurut (Long & Magerko, 2020a), literasi AI merupakan seperangkat kompetensi yang memungkinkan individu memahami, mengevaluasi, berkomunikasi, dan menggunakan teknologi AI secara efektif dan etis. Literasi AI tidak hanya mencakup kemampuan teknis dalam menggunakan aplikasi berbasis AI, tetapi juga mencakup pemahaman konsep dasar AI, kemampuan berpikir kritis terhadap output AI, kesadaran terhadap bias teknologi, serta pemahaman mengenai dampak sosial dan etika AI (Long & Magerko, 2020a). mengidentifikasi 17 kompetensi inti AI literacy, di antaranya *recognizing AI*, *data literacy*, *decision-making*, *machine learning steps*, *human roles in AI*, dan *ethics*. Konsep yang dikembangkan oleh Long dan Magerko tersebut menjadi salah satu kerangka literasi AI yang paling banyak digunakan dalam penelitian pendidikan dan pengembangan instrumen *AI literacy*.

Seiring meningkatnya perhatian terhadap pentingnya literasi AI, berbagai penelitian mulai mengembangkan instrumen pengukuran *AI literacy* pada siswa dan mahasiswa. Salah satu instrumen yang banyak dijadikan rujukan adalah instrumen literasi AI yang dikembangkan oleh Duri Long dan Brian Magerko. Instrumen tersebut dikembangkan berdasarkan kompetensi inti literasi AI yang meliputi pemahaman konsep AI, penggunaan AI, evaluasi sistem AI, etika AI, serta kesadaran kritis terhadap dampak teknologi AI. Selain itu, penelitian **Astrid Carolus** mengembangkan instrumen *MAILS (Meta AI Literacy Scale)* yang mengukur aspek *Use & Apply AI*, *Understand AI*, *Detect AI*, dan *AI Ethics* sebagai konstruk utama literasi AI. Meskipun instrumen tersebut telah digunakan dalam berbagai penelitian internasional, penerapannya pada konteks mahasiswa Indonesia, khususnya di Sulawesi Tenggara, masih memerlukan pengujian ulang untuk memastikan kesesuaian budaya, bahasa, dan karakteristik responden lokal.

Adaptasi instrumen dari konteks internasional ke konteks lokal merupakan langkah penting dalam penelitian psikometrik. Instrumen yang valid di suatu negara belum tentu memiliki validitas yang sama

ketika digunakan pada populasi dengan latar belakang budaya, sosial, dan pendidikan yang berbeda (Bartram et al., 2018). Mahasiswa di Sulawesi Tenggara memiliki karakteristik sosial, akses teknologi, pengalaman digital, dan lingkungan pendidikan yang berbeda dibandingkan dengan konteks pengembangan instrumen asli. Oleh karena itu, instrumen literasi AI yang diadaptasi dari Long dan Magerko perlu diuji kembali untuk mengetahui tingkat validitas dan reliabilitasnya ketika digunakan pada mahasiswa di Sulawesi Tenggara. Pendekatan adaptasi instrumen lintas budaya penting dilakukan untuk menjaga kesetaraan makna konstruk dan akurasi interpretasi hasil pengukuran pada populasi yang berbeda.

Pengujian validitas dan reliabilitas instrumen merupakan tahap penting dalam penelitian untuk memastikan bahwa instrumen mampu mengukur konstruk yang dimaksud secara akurat dan konsisten. Validitas mengacu pada sejauh mana suatu instrumen benar-benar mengukur konstruk yang hendak diukur, sedangkan reliabilitas berkaitan dengan konsistensi hasil pengukuran dan minimnya kesalahan pengukuran (Peeters & Harpe, 2020). Dalam penelitian ini, pengujian psikometrik dilakukan untuk mengevaluasi kualitas item, tingkat konsistensi internal, serta kesesuaian konstruk literasi AI pada mahasiswa. Selain itu, pengujian instrumen pada konteks lokal menjadi penting karena instrumen yang valid pada suatu populasi belum tentu memiliki tingkat validitas yang sama ketika diterapkan pada populasi dengan latar belakang budaya, sosial, dan pendidikan yang berbeda. Oleh karena itu, penelitian ini diharapkan dapat menghasilkan instrumen literasi AI yang layak digunakan dalam konteks pendidikan tinggi di Indonesia, khususnya di Sulawesi Tenggara.

Selain fokus pada pengujian instrumen, penelitian ini juga mengkaji perbedaan literasi AI berdasarkan gender. Perbedaan gender dalam penguasaan teknologi digital telah menjadi perhatian dalam berbagai penelitian sebelumnya. Beberapa studi menunjukkan bahwa laki-laki cenderung memiliki tingkat kepercayaan diri lebih tinggi dalam penggunaan teknologi, sedangkan perempuan lebih menonjol dalam aspek etika, kehati-hatian, dan tanggung jawab digital (Møgelvang et al., 2024) .Penelitian terbaru mengenai penggunaan AI dalam pembelajaran menunjukkan bahwa laki-laki cenderung menggunakan AI secara lebih eksploratif, sedangkan perempuan menunjukkan tingkat skeptisisme dan perhatian etika yang lebih tinggi terhadap penggunaan AI (Maxwell et al., 2025). Namun, hasil penelitian mengenai perbedaan gender dalam literasi AI masih menunjukkan temuan yang beragam dan belum konsisten. Oleh karena itu, analisis perbedaan literasi AI antara mahasiswa laki-laki dan perempuan di Sulawesi Tenggara menjadi penting untuk memberikan gambaran empiris mengenai kondisi literasi AI mahasiswa pada konteks lokal.

Penelitian ini menjadi penting karena belum banyak kajian di Indonesia yang mengadaptasi dan menguji instrumen literasi AI berbasis kerangka (Long & Magerko, 2020a)pada mahasiswa perguruan tinggi. Selain itu, penelitian ini juga diharapkan mampu memberikan kontribusi dalam pengembangan instrumen *AI literacy* yang sesuai dengan karakteristik mahasiswa Indonesia serta memberikan informasi mengenai kondisi literasi AI berdasarkan gender. Hasil penelitian dapat menjadi dasar bagi

perguruan tinggi dalam menyusun kebijakan, kurikulum, dan program pelatihan literasi AI guna mempersiapkan mahasiswa menghadapi era transformasi digital berbasis kecerdasan buatan.

METODE

Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain survei cross-sectional untuk menguji validitas dan reliabilitas instrumen literasi *Artificial Intelligence* (AI) yang diadaptasi dari instrumen yang dikembangkan oleh (Long & Magerko, 2020a) pada mahasiswa di Sulawesi Tenggara. Analisis psikometrik bertujuan untuk menilai karakteristik pengukuran suatu instrumen melalui estimasi kemampuan responden (*person ability*) dan tingkat kesulitan item (*item difficulty*) pada satu skala logit yang sama. Pendekatan ini dipilih karena Model Rasch mampu menghasilkan pengukuran yang bersifat objektif, independen terhadap sampel maupun item, serta memberikan informasi yang lebih komprehensif dibandingkan pendekatan Classical Test Theory (CTT) (Bond et al., 2021).

Partisipan

Partisipan penelitian terdiri atas mahasiswa aktif dari beberapa perguruan tinggi di Sulawesi Tenggara. Teknik pengambilan sampel menggunakan *purposive sampling* dengan kriteria: (1) mahasiswa aktif program sarjana; (2) pernah menggunakan aplikasi berbasis AI seperti ChatGPT, Google Gemini, atau aplikasi AI lainnya; dan (3) bersedia menjadi responden penelitian secara sukarela. Jumlah partisipan sebanyak 102 mahasiswa Sulawesi Tenggara. Ukuran sampel tersebut dianggap memadai untuk analisis Rasch karena mampu menghasilkan estimasi parameter item yang stabil dan reliabel.

Instrumen Penelitian

Instrumen penelitian merupakan hasil adaptasi dari instrumen literasi AI yang dikembangkan oleh (Long & Magerko, 2020a) Proses adaptasi instrumen dilakukan berdasarkan prinsip adaptasi lintas budaya yang meliputi proses penerjemahan, penyesuaian konteks budaya, validasi ahli, dan uji coba instrumen. Pendekatan adaptasi lintas budaya penting dilakukan karena instrumen yang dikembangkan pada konteks budaya tertentu belum tentu memiliki karakteristik pengukuran yang sama ketika digunakan pada populasi berbeda.

Instrumen terdiri atas 34 item pernyataan dengan tujuh dimensi utama, yaitu:

1. *AI Usage* (Penggunaan AI)
2. *AI Understanding* (Pemahaman AI),
3. *AI Recognition* (Pengenalan AI)
4. *AI Ethics* (Etika AI)
5. *AI Creation/Development* (Pengembangan AI)
6. AI Self-Efficacy
7. AI Self-Competency

Pengumpulan Data

Data dikumpulkan secara daring menggunakan Google Form selama semester genap tahun akademik 2025/2026. Sebelum pengisian instrumen, seluruh partisipan diberikan informasi mengenai tujuan penelitian, kerahasiaan data, serta hak partisipan untuk menghentikan partisipasi kapan saja. Seluruh partisipasi bersifat sukarela dan anonim.

Analisis Data

Data dianalisis menggunakan perangkat lunak WINSTEPS versi 4.7.0 dengan pendekatan Model Rasch. Analisis dilakukan melalui beberapa tahapan, yaitu:

1. Reliabilitas & Summary Statistics
2. Separation Index
3. Item Fit / Misfit Order
4. Item Statistics (Validitas & Tingkat Kesulitan)
5. Person Fit
6. Wright Map (dalam bentuk tabel teks)

Etika Penelitian

Penelitian ini dilaksanakan sesuai prinsip etika penelitian dan mengacu pada prinsip kerahasiaan, persetujuan partisipan, dan partisipasi sukarela sebagaimana diterapkan dalam penelitian psikometrik modern. Seluruh identitas responden dijaga kerahasiaannya dan data hanya digunakan untuk kepentingan akademik.

HASIL DAN PEMBAHASAN

Dokumen ini menyajikan hasil analisis Rasch yang disusun secara sistematis berdasarkan output WINSTEPS dari instrumen MAILS (34 item; 102 responden). Interpretasi disusun mengikuti urutan analisis Rasch.

Reliabilitas dan Summary Statistics

Tabel 1. Summary of measured person

SUMMARY OF 102 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	114.4	34.0	.63	.25	1.01	-.73	.99	-.76
SEM	2.0	.0	.13	.01	.08	.33	.08	.32
P.SD	20.6	.0	1.32	.11	.81	3.27	.80	3.20
S.SD	20.7	.0	1.33	.11	.81	3.29	.80	3.21
MAX.	169.0	34.0	6.20	1.01	3.98	8.38	4.00	7.64
MIN.	35.0	34.0	-5.90	.23	.06	-8.29	.06	-8.17

REAL RMSE	.31	TRUE SD	1.28	SEPARATION 4.21	Person RELIABILITY .95
MODEL RMSE	.27	TRUE SD	1.29	SEPARATION 4.71	Person RELIABILITY .96
S.E. OF Person MEAN = .13					

Person RAW SCORE-TO-MEASURE CORRELATION = .97
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .96 SEM = 4.25
STANDARDIZED (50 ITEM) RELIABILITY = .97

Tabel 2. Summary of measured item

SUMMARY OF 34 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	343.1	102.0	.00	.14	1.00	-.11	.99	-.19
SEM	5.8	.0	.11	.00	.05	.35	.05	.29
P.SD	33.5	.0	.62	.00	.27	2.00	.27	1.65
S.SD	34.0	.0	.63	.00	.28	2.03	.28	1.67
MAX.	396.0	102.0	1.64	.14	1.46	3.04	1.47	2.57
MIN.	253.0	102.0	-1.02	.13	.54	-3.92	.53	-3.46

REAL RMSE	.15	TRUE SD	.61	SEPARATION 4.16	Item	RELIABILITY .95
MODEL RMSE	.14	TRUE SD	.61	SEPARATION 4.40	Item	RELIABILITY .95
S.E. OF Item MEAN = .11						

Item RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000

Hasil analisis Rasch menunjukkan bahwa instrumen Modified AI Literacy Scale (MAILS) memiliki reliabilitas person dan reliabilitas item yang sangat tinggi, masing-masing sekitar 0,95, dengan nilai Cronbach Alpha sebesar 0,96–0,97. Temuan ini menunjukkan bahwa instrumen memiliki konsistensi internal yang sangat baik dalam mengukur konstruk literasi Artificial Intelligence (AI). Selain itu, nilai person separation sekitar 4,3 dan item separation sekitar 4,4 mengindikasikan bahwa instrumen mampu membedakan kemampuan responden dan tingkat kesulitan item secara jelas.

Menurut (Bond et al., 2021), reliabilitas person menggambarkan konsistensi pola respons responden, sedangkan reliabilitas item menunjukkan kestabilan hierarki tingkat kesulitan item apabila instrumen digunakan pada kelompok responden yang berbeda. Berbeda dengan Classical Test Theory (CTT) yang hanya menghasilkan satu koefisien reliabilitas, Model Rasch memberikan estimasi reliabilitas person dan item secara terpisah sehingga menghasilkan informasi yang lebih komprehensif mengenai kualitas instrumen.

Hasil penelitian ini sejalan dengan (Sumintono & Widhiarso, 2015) yang menjelaskan bahwa reliabilitas person dan reliabilitas item dalam Model Rasch menunjukkan konsistensi pengukuran serta kestabilan instrumen dalam mengukur konstruk yang sama. Nilai reliabilitas yang tinggi pada penelitian ini mengindikasikan bahwa instrumen Modified AI Literacy Scale (MAILS) mampu menghasilkan pengukuran yang konsisten sehingga layak digunakan sebagai alat ukur literasi AI mahasiswa.

Temuan ini juga mendukung konsep AI Literacy yang dikemukakan oleh (Ng et al., 2021) bahwa literasi AI merupakan konstruk multidimensional yang mencakup pengetahuan, keterampilan, pemahaman, kemampuan berpikir kritis, dan kesadaran etis dalam penggunaan AI. Oleh karena itu, instrumen yang digunakan untuk mengukur literasi AI perlu memiliki kualitas psikometrik yang baik agar setiap dimensi dapat diukur secara akurat dan konsisten.

Separation Index

Berdasarkan Summary of Measured Persons (Tabel1), diperoleh Person Separation sebesar 4,30 dengan reliabilitas person sebesar 0,95. Nilai tersebut menunjukkan bahwa instrumen mampu mengelompokkan responden ke dalam beberapa strata kemampuan literasi AI secara jelas. Selanjutnya, berdasarkan Summary of Measured Items, diperoleh Item Separation sebesar 4,40 dengan reliabilitas

item sebesar 0,95, yang menunjukkan bahwa tingkat kesulitan item tersebar dengan baik dan stabil pada skala logit..

Nilai person separation sebesar sekitar 4,3 menunjukkan bahwa instrumen mampu mengelompokkan mahasiswa ke dalam beberapa strata kemampuan literasi AI. Sementara itu, item separation sebesar sekitar 4,4 menunjukkan bahwa tingkat kesulitan item tersebar dengan baik sehingga mampu mengukur mahasiswa dengan kemampuan yang beragam.

Menurut (Bond et al., 2021), *person separation* dan *item separation* merupakan indikator penting dalam Model Rasch yang menunjukkan kemampuan instrumen membedakan tingkat kemampuan responden maupun tingkat kesulitan item. Semakin tinggi nilai separation, semakin baik kemampuan instrumen dalam menghasilkan pengukuran yang presisi. (Linacre, 2002) menjelaskan bahwa separation merupakan rasio antara variasi kemampuan yang sesungguhnya dengan kesalahan pengukuran sehingga nilai separation yang tinggi menunjukkan kualitas diskriminasi instrumen yang semakin baik.

Berdasarkan rumus strata Rasch yang dikemukakan oleh (Sumintono & Widhiarso, 2015),

$$H = \frac{(4 \times \text{Separation}) + 1}{3}$$

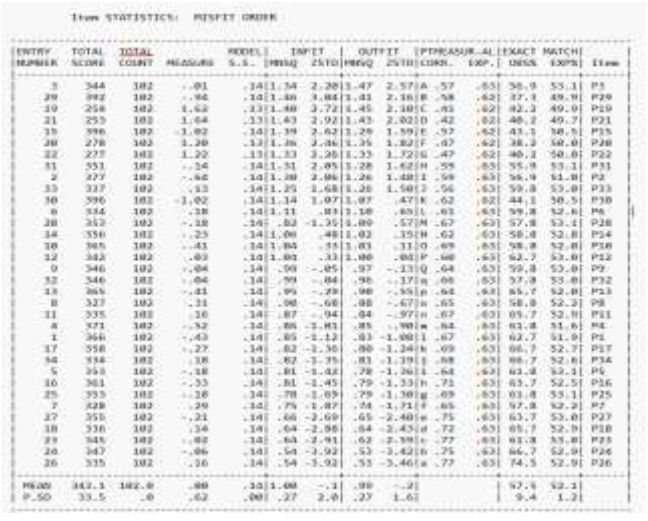
nilai person separation sebesar 4,30 menghasilkan:

$$H = \frac{(4 \times 4{,}30) + 1}{3} = \frac{18{,}2}{3} = 6{,}07$$

atau sekitar enam strata kemampuan. Hasil tersebut menunjukkan bahwa instrumen MAILS mampu mengelompokkan mahasiswa ke dalam enam kategori kemampuan literasi AI yang berbeda. Semakin banyak strata yang terbentuk menunjukkan bahwa instrumen memiliki sensitivitas yang tinggi dalam membedakan kemampuan responden sehingga mampu memberikan pengukuran yang lebih akurat

Item Fit / Misfit Order

Table 3. Item Statistics: Misfit Order



Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
3	344	102	-.01	.14	1.54	2.28	1.47	2.57	A .57	.65	56.9	55.1	P3
29	392	102	-.96	.16	1.46	1.84	1.41	2.16	B .58	.62	47.1	49.9	P29
19	356	102	1.62	.13	1.48	3.72	1.45	2.10	C .61	.62	42.3	49.0	P19
21	353	102	1.64	.13	1.43	3.92	1.43	2.02	D .62	.62	48.2	49.7	P21
35	396	102	-1.02	.14	1.39	2.62	1.29	1.59	E .57	.62	43.1	50.5	P35
20	370	102	1.30	.13	1.36	3.06	1.35	1.82	F .67	.62	38.2	50.0	P20
22	377	102	1.22	.13	1.33	3.26	1.33	1.72	G .67	.62	40.2	50.0	P22
31	351	102	-.14	.14	1.31	2.05	1.28	1.62	H .59	.65	55.9	53.1	P31
2	377	102	-.66	.14	1.30	2.06	1.26	1.48	I .59	.63	56.9	51.8	P2
33	337	102	.13	.14	1.25	1.68	1.26	1.50	J .56	.63	59.8	53.0	P33
30	396	102	-1.02	.14	1.14	1.07	1.07	.47	K .62	.62	44.1	50.5	P30
6	334	102	.18	.14	1.11	.83	1.10	.69	L .63	.65	59.8	52.6	P6
28	353	102	-.18	.14	.82	-1.95	1.00	.57	M .67	.63	57.8	53.1	P28
14	350	102	-.25	.14	1.00	.48	1.02	.35	N .62	.63	50.0	52.8	P14
18	365	102	-.41	.14	1.04	.55	1.03	.31	O .69	.65	58.8	52.8	P18
12	342	102	.03	.14	1.03	.33	1.00	.06	P .68	.63	62.7	53.0	P12
9	346	102	-.04	.14	.90	-.85	.97	-.13	Q .64	.63	59.8	53.0	P9
32	346	102	-.04	.14	.89	-.94	.96	-.17	R .66	.63	57.8	53.0	P32
13	365	102	-.41	.14	.96	-.29	.90	-.55	S .64	.63	65.7	52.8	P13
8	327	102	.31	.14	.98	-.68	.88	-.67	T .65	.63	58.8	52.3	P8
11	335	102	.16	.14	.87	-.94	.84	-.97	U .67	.63	65.7	52.9	P11
4	371	102	-.52	.14	.86	-1.83	.85	-.98	V .64	.63	63.8	51.6	P4
1	366	102	-.43	.14	.85	-2.12	.81	-1.08	W .67	.63	62.7	51.9	P1
17	358	102	-.27	.14	.82	-2.36	.80	-1.24	X .69	.63	66.7	52.7	P17
34	334	102	.18	.14	.82	-3.35	.83	-1.39	Y .68	.65	68.7	52.6	P34
5	353	102	-.18	.14	.81	-3.43	.78	-1.36	Z .68	.63	63.8	53.1	P5
16	361	102	-.33	.14	.81	-3.45	.79	-1.33	a .71	.63	63.7	52.5	P16
25	355	102	-.18	.14	.78	-3.69	.79	-1.30	g .69	.63	63.8	53.1	P25
7	328	102	.29	.14	.75	-3.87	.74	-1.71	f .65	.63	57.8	52.2	P7
27	355	102	-.21	.14	.66	-2.69	.65	-2.40	e .75	.63	63.7	53.0	P27
18	336	102	.14	.14	.64	-2.98	.64	-2.45	d .72	.63	65.7	52.9	P18
23	345	102	-.02	.14	.64	-2.91	.62	-2.59	c .77	.63	61.8	53.0	P23
26	347	102	-.06	.14	.54	-3.92	.53	-3.42	b .75	.63	66.7	52.9	P26
26	335	102	.16	.14	.54	-3.92	.53	-3.46	a .77	.63	74.5	52.9	P26
MEAN	342.1	102.0	.00	.13	1.00	-.1	.99	-.2			57.6	52.1	
P.SD	33.5	.0	.62	.00	.27	2.0	.27	1.6			9.4	1.2	

Tabel 4. Hasil Analisis Kesesuaian Item (Item Fit Statistics) Instrumen MAILS Berdasarkan Model Rasch

No.	Kode Item	Dimensi	Pernyataan Instrumen MAILS	Measure (Logit)	Infit MNSQ	Outfit MNSQ	Pt-Measure Corr.	Kategori
1	AIL1	Penggunaan	Saya mampu mengoperasikan aplikasi AI dalam kehidupan sehari-hari	-0.43	0.85	0.83	0.67	Fit
2	AIL2	Penggunaan	Saya mampu menggunakan aplikasi AI untuk mempermudah aktivitas sehari-hari saya	-0.64	1.30	1.26	0.59	Borderline
3	AIL3	Penggunaan	Saya mampu menggunakan aplikasi AI secara efektif untuk mencapai tujuan harian saya	-0.01	1.34	1.47	0.57	Misfit
4	AIL4	Penggunaan	Saya mampu berinteraksi dengan aplikasi AI untuk mempermudah tugas-tugas saya dalam kehidupan sehari-hari	-0.52	0.86	0.85	0.64	Fit
5	AIL5	Penggunaan	Saya mampu bekerja secara efektif bersama aplikasi berbasis teknologi AI	-0.18	0.81	0.78	0.64	Fit
6	AIL6	Penggunaan	Saya mampu berkomunikasi dengan aplikasi AI secara efektif dalam aktivitas sehari-hari	0.18	1.11	1.10	0.61	Fit
7	AIL7	Pemahaman	Saya mengetahui konsep-konsep paling penting dari topik yang berkaitan dengan teknologi AI	0.29	0.75	0.74	0.65	Fit
8	AIL8	Pemahaman	Saya mengetahui definisi dasar dari AI	0.31	0.90	0.88	0.65	Fit

9	AIL9	Pemahaman	Saya mampu menilai keterbatasan dan peluang yang ditawarkan oleh aplikasi AI	-0.04	0.99	0.97	0.64	Fit
10	AIL10	Pemahaman	Saya mampu menilai keuntungan dan kerugian dari penggunaan aplikasi AI	-0.41	1.04	1.01	0.69	Fit
11	AIL11	Pemahaman	Saya mampu memikirkan kegunaan baru yang potensial dari teknologi AI	0.16	0.87	0.84	0.67	Fit
12	AIL12	Pemahaman	Saya mampu membayangkan kemungkinan penerapan dari teknologi AI di masa depan	0.03	1.04	1.00	0.60	Fit
13	AIL13	Pengenalan	Saya mampu mengetahui apakah aplikasi yang sedang saya gunakan menggunakan teknologi AI	-0.41	0.95	0.90	0.64	Fit
14	AIL14	Pengenalan	Saya mampu membedakan antara perangkat yang menggunakan AI dengan yang tidak	-0.23	1.06	1.02	0.62	Fit
15	AIL15	Pengenalan	Saya mampu membedakan apakah saya sedang berinteraksi dengan AI atau manusia nyata	-1.02	1.39	1.29	0.46	Borderline
16	AIL16	Etika	Saya mampu mempertimbangkan konsekuensi dari penggunaan teknologi AI bagi masyarakat	-0.33	0.81	0.79	0.71	Fit
17	AIL17	Etika	Saya memperhatikan aspek etika saat memutuskan	-0.27	0.82	0.80	0.69	Fit

			untuk menggunakan data dari teknologi AI					
18	AIL18	Etika	Saya mampu menganalisis aplikasi AI berdasarkan implikasi etikanya	0.14	0.64	0.64	0.72	Fit
19	CAI1	Pengembangan AI	Saya memiliki kemampuan untuk merancang aplikasi AI yang baru	1.62	1.40	1.45	0.41	Misfit
20	CAI2	Pengembangan AI	Saya memiliki kemampuan untuk memprogram aplikasi baru di bidang AI	1.20	1.36	1.35	0.47	Borderline
21	CAI3	Pengembangan AI	Saya memiliki kemampuan untuk mengembangkan aplikasi AI baru yang inovatif	1.64	1.43	1.43	0.42	Misfit
22	CAI4	Pengembangan AI	Saya mampu memilih tools yang tepat (misalnya framework, IDE, SDK, bahasa pemrograman) untuk memprogram aplikasi AI	1.22	1.33	1.33	0.47	Borderline
23	AISE1	Self-Efficacy	Saya dapat mengandalkan keterampilan saya dalam situasi sulit saat menggunakan AI	-0.02	0.64	0.62	0.77	Fit
24	AISE2	Self-Efficacy	Saya dapat menangani sebagian besar masalah terkait AI dengan baik secara mandiri	-0.06	0.54	0.53	0.75	Fit
25	AISE3	Self-Efficacy	Saya dapat menyelesaikan tugas yang rumit dan menantang	-0.18	0.78	0.79	0.69	Fit

			saat bekerja dengan AI					
26	AISE4	Self-Efficacy	Saya selalu mengikuti inovasi terbaru yang berkaitan dengan aplikasi AI	0.16	0.54	0.53	0.77	Fit
27	AISE5	Self-Efficacy	Meskipun perubahan yang terjadi di bidang AI sangat cepat, namun saya selalu bisa meng-update pengetahuan saya	-0.21	0.66	0.65	0.75	Fit
28	AISE6	Self-Efficacy	Saya berhasil tetap untuk terus up-to-date meskipun banyak aplikasi AI baru yang terus bermunculan saat ini	-0.18	0.82	1.09	0.67	Fit
29	AISC1	Self-Competency	Saya mampu untuk tidak membiarkan AI mempengaruhi keputusan yang akan saya buat dalam kehidupan sehari-hari saya	1.62	1.46	1.41	0.41	Misfit
30	AISC2	Self-Competency	Saya mampu mencegah AI untuk mempengaruhi keputusan penting dalam hidup saya	-1.02	1.14	1.07	0.62	Fit
31	AISC3	Self-Competency	Saya bisa menyadari jika teknologi AI telah mempengaruhi keputusan yang saya buat dalam kehidupan sehari-hari saya	-0.14	1.31	1.28	0.59	Borderline
32	AISC4	Self-Competency	Saya mampu mengendalikan perasaan frustrasi atau	-0.04	0.99	0.96	0.66	Fit

			kecemasan yang muncul saat sedang menggunakan teknologi AI					
33	AISC5	Self-Competency	Saya mampu mengatasi rasa frustrasi atau ketakutan yang muncul saat saya sedang berinteraksi dengan teknologi AI	0.13	1.25	1.26	0.56	Fit
34	AISC6	Self-Competency	Saya mampu mengontrol rasa euforia yang muncul saat saya berhasil menyelesaikan tugas saya dengan menggunakan aplikasi AI	0.18	0.82	0.81	0.68	Fit

Berdasarkan tabel Item Statistics: Misfit Order, sebagian besar item memenuhi kriteria kecocokan model Rasch. Item P3, P19, P21 dan P29 memiliki nilai Outfit MNSQ relatif tinggi sehingga memerlukan peninjauan lebih lanjut. Namun seluruh item memiliki nilai Point Measure Correlation positif sehingga tetap mengukur konstruk yang sama.

Sebagian besar item memenuhi kriteria fit terhadap Model Rasch. Dari 34 item, sebanyak 25 item berada pada kategori fit, 5 item berada pada kategori borderline, sedangkan 4 item tergolong misfit, yaitu AIL3, CAI1, CAI3, dan AISC1.

Item dinyatakan sesuai dengan Model Rasch apabila memiliki nilai Outfit Mean Square (MNSQ) antara 0,50–1,50, karena rentang tersebut masih dianggap produktif untuk pengukuran (Linacre, 2002). Statistik MNSQ digunakan untuk mengevaluasi tingkat kesesuaian antara pola respons empiris dengan ekspektasi Model Rasch. Nilai yang mendekati 1,00 menunjukkan kesesuaian yang baik, sedangkan penyimpangan yang besar dari nilai tersebut mengindikasikan adanya *misfit* yang dapat memengaruhi kualitas pengukuran (Bond et al., 2021).

Item AIL3 yang mengukur kemampuan menggunakan AI secara efektif untuk mencapai tujuan sehari-hari menunjukkan variasi respons yang relatif tinggi. Hal ini kemungkinan disebabkan oleh perbedaan intensitas penggunaan AI di kalangan mahasiswa. Sebagian mahasiswa telah memanfaatkan AI sebagai alat pendukung belajar secara optimal, sedangkan sebagian lainnya masih menggunakan AI hanya sebagai mesin pencari informasi.

Temuan tersebut mendukung konsep AI Use Competency yang dikemukakan oleh (Long & Magerko, 2020b), bahwa kemampuan menggunakan AI secara efektif bukan hanya berkaitan dengan kemampuan teknis mengoperasikan aplikasi AI, tetapi juga kemampuan mengintegrasikan AI dalam penyelesaian masalah secara produktif.

Item CAI1 dan CAI3 yang berkaitan dengan kemampuan mengembangkan aplikasi AI juga menunjukkan kategori misfit. Kondisi tersebut sangat mungkin disebabkan oleh karakteristik responden yang sebagian besar bukan berasal dari program studi informatika sehingga pengalaman mereka dalam merancang maupun memprogram sistem AI masih sangat terbatas. Oleh karena itu, variasi respons lebih mencerminkan heterogenitas pengalaman responden dibandingkan kelemahan konstruk instrumen.
(Boone et al., 2014) menjelaskan bahwa item misfit tidak harus langsung dihapus karena masih mungkin memiliki relevansi teoritis terhadap konstruk yang diukur. Revisi redaksional atau penyesuaian konteks item sering kali menjadi pilihan yang lebih tepat dibandingkan eliminasi item.

Item Statistics (Validitas dan Tingkat Kesulitan)

Tabel 5. Item Statistics Instrumen Modified AI Literacy Scale (MAILS) Berdasarkan Model Rasch

Kategori	Kode Item	Jumlah
Fit	AIL1, AIL4, AIL5, AIL6, AIL7, AIL8, AIL9, AIL10, AIL11, AIL12, AIL13, AIL14, AIL16, AIL17, AIL18, AISE1, AISE2, AISE3, AISE4, AISE5, AISE6, AISC2, AISC4, AISC5, AISC6	**25**
Borderline	AIL2, AIL15, CAI2, CAI4, AISC3	**5**
Misfit	AIL3, CAI1, CAI3, AISC1	**4**

Tabel 6. Tingkat Kesulitan Item Berdasarkan Nilai Logit

Kategori	Rentang Logit	Item
Sangat Sulit	> 1,00	CAI3 (1,64), CAI1 (1,62), CAI4 (1,22), CAI2 (1,20), AISC1 (1,62)
Sedang	-1,00 s.d. 1,00	Sebagian besar item (AIL1–AIL14, AIL16–AIL18, AISE1–AISE6, AISC3–AISC6)
Sangat Mudah	< -1,00	AIL15 (-1,02), AISC2 (-1,02)

Hasil analisis Item Statistics menunjukkan bahwa dari 34 item instrumen Modified AI Literacy Scale (MAILS), sebanyak 25 item (73,53%) memenuhi kategori fit, 5 item (14,71%) berada pada kategori borderline, dan 4 item (11,76%) tergolong misfit. Klasifikasi tersebut didasarkan pada nilai Infit Mean Square (MNSQ) dan Outfit Mean Square (MNSQ) yang digunakan sebagai indikator kesesuaian item terhadap Model Rasch. Menurut Linacre (2024), item dikatakan sesuai dengan Model Rasch apabila memiliki nilai MNSQ pada kisaran 0,50–1,50, sedangkan nilai yang mendekati 1,00 menunjukkan tingkat kecocokan yang paling ideal.

Sebagian besar item menunjukkan nilai Infit dan Outfit MNSQ yang mendekati satu, sehingga dapat disimpulkan bahwa item-item tersebut telah bekerja sesuai dengan ekspektasi Model Rasch dan mampu mengukur konstruk literasi AI secara konsisten. Sementara itu, item AIL3, CAI1, CAI3, dan AISC1 menunjukkan nilai Outfit MNSQ relatif lebih tinggi dibandingkan item lainnya sehingga

dikategorikan sebagai misfit. Kondisi ini mengindikasikan bahwa respons responden terhadap item tersebut lebih bervariasi daripada yang diprediksi oleh model. Menurut (Boone et al., 2014), item yang tergolong misfit tidak harus langsung dieliminasi, tetapi perlu ditelaah kembali dari sisi substansi, redaksi, maupun kesesuaiannya dengan karakteristik responden.

Berdasarkan nilai Measure (logit), tingkat kesulitan item berada pada rentang −1,02 hingga 1,64 logit, yang menunjukkan bahwa instrumen memiliki variasi tingkat kesulitan yang cukup luas. Item CAI3 (1,64 logit), CAI1 (1,62 logit), AISC1 (1,62 logit), CAI4 (1,22 logit), dan CAI2 (1,20 logit) merupakan item dengan tingkat kesulitan tertinggi. Seluruh item tersebut berasal dari dimensi Pengembangan AI dan Self-Competency, yang menuntut kemampuan lebih kompleks seperti mengembangkan aplikasi AI, memilih perangkat pengembangan, serta mempertahankan kendali terhadap keputusan yang dipengaruhi AI.

Sebaliknya, item AIL15 dan AISC2 memiliki nilai logit −1,02, sehingga termasuk kategori sangat mudah. Kedua item tersebut lebih mudah disetujui oleh responden karena berkaitan dengan kemampuan mengenali interaksi dengan AI serta mempertahankan keputusan penting dari pengaruh AI. Menurut (Bond et al., 2021) , distribusi nilai logit yang luas menunjukkan bahwa instrumen memiliki measurement coverage yang baik karena mampu mengukur responden pada berbagai tingkat kemampuan.

Secara keseluruhan, hasil analisis Item Statistics menunjukkan bahwa instrumen MAILS memiliki kualitas psikometrik yang baik. Mayoritas item memenuhi kriteria kecocokan Model Rasch dan memiliki tingkat kesulitan yang bervariasi, sehingga instrumen mampu mengukur literasi AI mahasiswa secara komprehensif. Meskipun terdapat beberapa item yang tergolong borderline dan misfit, item-item tersebut masih dapat dipertahankan dengan melakukan revisi redaksional atau evaluasi substansi pada penelitian selanjutnya.

Person Fit

Tabel 7. Responden dengan Pola Respons Tidak Sesuai (Most Unexpected Responses)

No.	Kode Responden	Kode Item	Skor Observasi	Skor Ekspektasi	Residual	Interpretasi
1	062P	AIL3	1	4.62	-3.62	Responden dengan kemampuan tinggi memberikan skor jauh lebih rendah dari yang diprediksi model.
2	099P	AISE6	1	1.03	0.97	Respons berbeda dari ekspektasi model, namun selisih relatif kecil.
3	072P	AISC5	1	4.03	-3.03	Respons menunjukkan ketidakkonsistenan terhadap estimasi kemampuan.
4	081P	CAI2	1	3.91	-2.91	Respons lebih rendah dibandingkan prediksi model.

5	028P	AIL15	5	1.99	3.01	Responden dengan kemampuan rendah memberikan skor jauh lebih tinggi dari prediksi model.

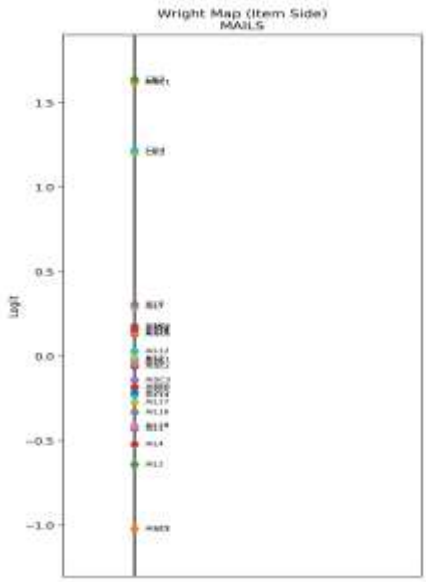
Catatan: Tabel di atas hanya menampilkan beberapa contoh *Most Unexpected Responses*. Analisis lengkap terdapat pada output Winsteps.

Analisis Most Unexpected Responses menunjukkan bahwa hanya sebagian kecil responden memiliki pola jawaban yang tidak sesuai dengan ekspektasi Model Rasch. Kondisi tersebut diduga disebabkan oleh ketidakkonsistenan dalam menjawab, kurang memahami pernyataan, atau pengalaman penggunaan AI yang berbeda.

Person misfit menunjukkan adanya pola respons yang tidak sesuai dengan ekspektasi model. Pola tersebut dapat muncul karena berbagai faktor, seperti menebak jawaban, kelelahan, kurangnya perhatian, atau strategi respons tertentu, sehingga tidak selalu mencerminkan rendahnya kualitas instrument (Dünya, 2024).

Wright Map

Wright Map digunakan untuk membandingkan distribusi kemampuan responden dengan distribusi kesulitan item. Secara umum penyebaran item telah mencakup rentang kemampuan responden, meskipun terdapat beberapa item yang sangat mudah dan sangat sulit. Hal ini menunjukkan instrumen telah memiliki targeting yang cukup baik, namun penambahan item pada tingkat kesulitan sedang masih dapat meningkatkan presisi pengukuran.



Grafik ini menyajikan distribusi tingkat kesulitan item berdasarkan nilai logit. Item dengan logit tertinggi (CAI3, CAI1, AISC1) merupakan item tersulit, sedangkan AIL15 dan AISC2 merupakan item termudah.

Hasil Wright Map menunjukkan bahwa distribusi tingkat kesulitan item telah mencakup sebagian besar rentang kemampuan mahasiswa. Meskipun demikian, masih terdapat beberapa celah pada

tingkat kemampuan sedang sehingga penambahan item dengan tingkat kesulitan menengah dapat meningkatkan ketepatan pengukuran.

Wright Map (*person–item map*) merupakan visualisasi yang menempatkan kemampuan responden dan tingkat kesulitan item pada skala logit yang sama. Melalui pemetaan ini, peneliti dapat mengevaluasi kesesuaian (*targeting*) antara distribusi kemampuan responden dan tingkat kesulitan item, serta mengidentifikasi area konstruk yang belum terwakili secara memadai oleh instrument (Shahat et al., 2024)

KESIMPULAN

Hasil analisis menggunakan Model Rasch menunjukkan bahwa Modified AI Literacy Scale (MAILS) memiliki kualitas psikometrik yang baik untuk mengukur literasi Artificial Intelligence (AI) mahasiswa. Instrumen menunjukkan reliabilitas yang sangat tinggi serta mampu membedakan kemampuan responden dan tingkat kesulitan item secara baik. Sebagian besar item memenuhi kriteria fit, sedangkan beberapa item yang berada pada kategori *borderline* dan *misfit* memerlukan penyempurnaan redaksional tanpa harus dieliminasi. Variasi tingkat kesulitan item menunjukkan bahwa instrumen mampu mengukur literasi AI pada berbagai tingkat kemampuan mahasiswa. Dengan demikian, MAILS layak digunakan sebagai instrumen pengukuran literasi AI di pendidikan tinggi Indonesia serta berpotensi mendukung penelitian dan evaluasi pembelajaran berbasis AI di masa mendatang.

DAFTAR PUSTAKA

Bartram, D., Berberoglu, G., Grégoire, J., Hambleton, R., Muniz, J., & van de Vijver, F. (2018). ITC Guidelines for Translating and Adapting Tests (Second Edition). *International Journal of Testing*, *18*(2), 101–134. https://doi.org/10.1080/15305058.2017.1398166

Bond, T. G., Yan, Z., & Heene, M. (2021). Applying the Rasch Model. In *Applying the Rasch Model*. Routledge. https://doi.org/10.4324/9781410614575

Boone, W. J., Yale, M. S., & Staver, J. R. (2014). Rasch analysis in the human sciences. In *Rasch Analysis in the Human Sciences*. https://doi.org/10.1007/978-94-007-6857-4

Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS - Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. *Computers in Human Behavior: Artificial Humans*, *1*(2), 100014. https://doi.org/10.1016/j.chbah.2023.100014

Dünya, B. A. (2024). EDITORIAL: Adapting the Person Fit Analysis: Ideas on Detecting Person Misfit in Computerized Adaptive Testing. *Journal of Measurement and Evaluation in Education and Psychology*, *15*(1), 1–4. https://doi.org/10.21031/epod.1461703

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer,

J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*(March). https://doi.org/10.1016/j.lindif.2023.102274

Linacre, J. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Meas Trans*, *16*.

Long, D., & Magerko, B. (2020a). What is AI Literacy? Competencies and Design Considerations. *Conference on Human Factors in Computing Systems - Proceedings*, 1–16. https://doi.org/10.1145/3313831.3376727

Long, D., & Magerko, B. (2020b). What is AI Literacy? Competencies and Design Considerations. *Conference on Human Factors in Computing Systems - Proceedings*, 1–16. https://doi.org/10.1145/3313831.3376727

Maxwell, D., Oyarzun, B., Kim, S., & Bong, J. Y. (2025). Generative AI in Higher Education: Demographic Differences in Student Perceived Readiness, Benefits, and Challenges. *TechTrends*, *69*(6), 1248–1259. https://doi.org/10.1007/s11528-025-01109-6

Møgelvang, A., Bjelland, C., Grassini, S., & Ludvigsen, K. (2024). Gender Differences in the Use of Generative Artificial Intelligence Chatbots in Higher Education: Characteristics and Consequences. *Education Sciences*, *14*(12), 1–19. https://doi.org/10.3390/educsci14121363

Ng, D. T. K., Chu, S., Shen, M., & Leung, J. (2021). *AI Literacy: Definition, Teaching, Evaluation and Ethical Issues*.

Peeters, M. J., & Harpe, S. E. (2020). Updating conceptions of validity and reliability. *Research in Social and Administrative Pharmacy*, *16*(8), 1127–1130. https://doi.org/https://doi.org/10.1016/j.sapharm.2019.11.017

Shahat, M. A., Boone, W. J., Al-Alawi, K. A., & Al-Balushi, S. M. (2024). Rasch analysis in developing a questionnaire to measure self-efficacy beliefs of Omani preservice science teachers for teaching through engineering design. *Humanities and Social Sciences Communications*, *11*(1), 1–17. https://doi.org/10.1057/s41599-024-04087-x

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*.

Unesco. (2024). AI competency framework for teachers. *AI Competency Framework for Teachers*. https://doi.org/10.54675/zjte2084