BenchCouncil Transactions on Benchmarks, Standards and Evaluations 5 (2025) 100254

Contents lists available at ScienceDirect

BenchCouncil Transactions on Benchmarks, Standards and Evaluations

journal homepage: www.keaipublishing.com/en/journals/benchcouncil-transactions-onbenchmarks-standards-and-evaluations/

Full Length Article

"We don't plagiarise, we parrot": Cognitive load and ethical perceptions in higher education written assessment

Ibnatul Jalilah Yusof *, Zakiah Mohamad Ashari, Lukman Hakim Ismail, Mira Panadi

Universiti Teknologi Malaysia, Malaysia

A R T I C L E I N F O

Keywords:
Cognitive load
Ethical perceptions
Parroting behaviour
Plagiarisms

A B S T R A C T

Generative artificial intelligence has reshaped written assessment in higher education and sharpened concerns about "parroting," the undisclosed use of AI-generated text with minimal cognitive engagement. This study examines the cognitive and ethical mechanisms underlying parroting among undergraduates in one Malaysian research university. Drawing on Cognitive Load Theory and Dual-System Theory, parroting is conceptualised across three dimensions: intrinsic load, extraneous load, and ethical rationalisation. Survey responses from 211 students were analysed using Rasch measurement to evaluate item reliability, construct separation, differential item functioning (DIF) across academic fields and item hierarchies. Results indicate that items function equivalently for engineering, non-engineering, and science students, supporting the instrument's fairness and stability. Overall, findings show that parroting is most strongly driven by extraneous pressures such as vague instructions and heavy workload, followed by intrinsic challenges related to writing confidence and conceptual understanding. Ethical rationalisation is endorsed least frequently but becomes more salient when institutional guidance on AI use is unclear. The study offers implications for pedagogy and policy, underscoring the need for explicit AI-use guidelines, improved task design, and learning environments that promote ethically responsible engagement with generative technologies.

Background

The advent of generative artificial intelligence (GenAI) tools such as ChatGPT has profoundly transformed perspectives on written assessments in higher education [1,2]. These tools offer students instant access to grammatically polished, coherent text with minimal cognitive effort. While this may appear to support productivity, a more insidious challenge has emerged: the increasing prevalence of cognitively superficial writing, whereby students submit polished work that lacks evidence of internalisation, reflection, or intellectual authorship [3–5].

This phenomenon aligns with what scholars refer to as *parroting*; the act of reproducing content verbatim or with minimal alterations, devoid of meaningful comprehension or engagement [6,7]. Traditionally, parroting has been associated with rote learning and low-level memorisation [8,9], but in the context of AI-generated content, it signals a more complex and ethically ambiguous behaviour. The student's work may appear fluent and structured, yet lacks the deeper cognitive investment expected in higher education writing tasks.

Why do students do it? Why do they submit assignments that are not the product of their own thinking, but rather stitched together from AI-generated text? This study assumes that such behaviour is not simply a matter of laziness or rebellion, but a reflection of deeper psychological and situational mechanisms.

Firstly, consider the existing definitions of plagiarism. Whether framed in legal, institutional, or scholarly terms, plagiarism is typically defined as the use of another human's intellectual work without proper attribution [10,11]. However, AI-generated text does not originate from a person, and current plagiarism frameworks are often ill-equipped to address the complexities introduced by AI-assisted writing [12,13]. While the Committee on Publication Ethics [14] maintains that AI cannot be listed as an author due to its lack of agency and accountability, many students may continue to use such tools in ways that obscure their own contribution.

This grey area has led to a dangerous misconception; that parroting AI-generated content is acceptable because *'it is not technically plagiarism'*. Furthermore, as institutions race to integrate AI literacy into curricula, few have developed clear policies addressing the boundaries between AI support and AI substitution [1]. In the absence of clear

Peer review under the responsibility of The International Open Benchmark Council.

* Corresponding author.
E-mail address: ijalilah@utm.my (I.J. Yusof).

https://doi.org/10.1016/j.tbench.2025.100254
Received 25 August 2025; Received in revised form 6 December 2025; Accepted 10 December 2025
Available online 12 December 2025

ethical guidelines, students may normalise parroting as a legitimate academic strategy [1,4,15], especially under cognitive strain, time pressure, or a desire for efficiency.

Secondly, a more complex reason lies in the psychological interplay between Dual-System Theory (DST) and Cognitive Load Theory (CLT). Students may not always make deliberate ethical judgments; rather, they respond impulsively under pressure. For example, some may think, *"ChatGPT writes better than I ever could, why should I bother?"* is a reflection of System 1 thinking, the fast, intuitive, and automatic mode described in DST [16]. Others may feel cognitively overwhelmed by the task itself and conclude, *"This is too hard, AI can do it faster and better than me"* is a coping response aligned with CLT [17,18].

In both cases, students opt out of the mental effort required for authentic academic work, either because the task exceeds their working memory capacity especially in academic writing tasks requiring synthesis, originality, and technical precision, [19–21] or because an easier alternative presents itself [3–5]. This convergence of automatic decision-making and cognitive strain may explain why students rationalise GenAI use without engaging in meaningful writing processes.

Problem statement

Despite growing attention to the academic implications of GenAI, the specific phenomenon of parroting remains under-theorized and insufficiently measured, particularly within Malaysia higher education [22]. As such, parroting exists in a grey zone technically legal, ethically ambiguous, and psychologically convenient. Yet, little is known about how students themselves rationalise or justify this behaviour, and under what conditions they are more likely to engage in it.

Therefore, the present study focuses on two key areas:

(i) the cognitive (intrinsic and extraneous load) mechanisms that lead students to rely on AI-generated content in their academic writing,
(ii) their ethical perceptions regarding the acceptability of such practices.

Literature review

Parroting is not a new term in educational discourse. It draws from the observable behaviour of parrots mimicking sounds without true comprehension. While animal cognition research continues to debate the cognitive capacities of parrots, particularly in the case of *Alex the African Grey* who demonstrated the ability to respond meaningfully to prompts [23], this study adopts a more conservative interpretation. Here, parroting refers to surface-level mimicry that lacks cognitive depth, shaped more by repetition and performance than by meaningful understanding.

In classroom contexts, parroting has traditionally been associated with rote memorisation, where students reproduce information with little or no conceptual engagement [8,9]. Such practices may yield correct answers but rarely reflect deep learning or analytical thought. The concern becomes even more pressing when parroting migrates into written assessments, where the stakes are higher, and the appearance of originality can mask a lack of authentic intellectual contribution. It takes on a more deliberate and strategic form.

Glatt [7] examined this phenomenon in students' written work and identified three interrelated drivers:

(i) Insecurity about writing ability, which leads learners to appropriate existing text rather than risk producing original prose; and
(ii) Time pressure, which encourages shortcuts that prioritize rapid completion over thoughtful composition; and
(iii) Low motivation, prompting students to avoid the effortful stages of planning, drafting, and revision.

Glatt's [7] conceptualization of parroting reflects a traditional academic context, where students, despite resorting to copying, would still engage with the source material by reading and minimally modifying it. Parroting was thus a low-effort strategy, but not entirely void of cognitive involvement. In contrast, the rise of GenAI introduces a more detached form of mimicry, one that allows students to bypass even the most basic interaction with the text. Building on this, the present study conceptualizes parroting not merely as mindless copying but as a form of cognitive outsourcing, where students offload the demanding aspects of intellectual work to an external system. This shift invites deeper inquiry into how such behaviour is shaped not only by academic pressure but also by the cognitive architecture of decision making and the ethical reasoning (or absence thereof) that justifies it.

According to DST [16], human thinking operates through two modes which are System 1, which is fast, automatic, and intuitive, and System 2, which is slow, effortful, and reflective. In academic settings, writing typically demands System 2 processing such as analysing ideas, structuring arguments, and evaluating sources. However, when confronted with cognitively overwhelming tasks or vague assignment instructions, students often default to shortcuts such as copy-and-paste (System 1) [3,7,24]. This reliance on intuition over analysis is not a sign of laziness but a systematic response to cognitive strain, particularly when combined with high intrinsic or extraneous cognitive load [18].

Here, CLT complements DST. Writing tasks that involve abstract or unfamiliar content increase intrinsic load, while unclear instructions add extraneous load. Both can easily exceed students' working memory capacity. Instead of engaging in laborious thought, students may rely on GenAI tools as cognitive crutches [25] inserting or modifying text with minimal intellectual contribution. The ethical dimension deepens this issue. Students may engage in what Bandura [26] calls moral disengagement, a psychological process in which individuals justify unethical behaviour through diffusion of responsibility, minimization of harm, or ambiguity in norms [27,28]. If there are no explicit university policies about GenAI citation [1,4,15,29], or if peers do it without consequence [30,31]. In doing so, they enter a grey zone where academic mimicry masquerades as competence.

This tendency was operationalised in the present study, where parroting behaviours were measured across three dimensions:

(i) Intrinsic Load-Induced Parroting: When students perceive the task as beyond their own ability (e.g., abstract topic, unfamiliar concepts, requirement for synthesis), students may experience intrinsic cognitive overload. Believing that AI will produce better results than they could, they bypass deep thinking (System 2) and uncritically use AI-generated content.
(ii) Extraneous Load-Induced Parroting: When the assignment is poorly structured, vague, or confusing, students may encounter extraneous cognitive load. Rather than seeking clarification or attempting to navigate unclear instructions, they impulsively rely on AI as a shortcut, again engaging System 1 thinking.
(iii) Ethically Rationalised Parroting: Even when aware of academic norms, students may morally disengage, justifying their use of AI based on peer behaviour, lack of institutional guidelines, or the perception that AI-generated text is not 'real plagiarism'. This rationalisation supports System 1 decisions that minimize effort and suppress ethical reflection.

Methodology

Research design

This study employed a quantitative, cross-sectional survey design to investigate parroting behaviour defined as the undisclosed use of generative AI content with minimal cognitive engagement among education students. Drawing on Cognitive Load Theory and Dual-System Theory, parroting was conceptualized across three dimensions:

intrinsic load–induced, extraneous load–induced, and ethically rationalised behaviours. A Rasch measurement approach was adopted to examine the psychometric properties of the instrument and to establish endorsement hierarchies within each dimension.

Sample

Table 1 presents the distribution of respondents selected through purposive sampling. A total of 211 participants from one research university in Malaysia were included in this study. The largest proportion came from Non-Engineering/Social Sciences fields (n = 87, 41.2%), representing disciplines such as Education, Business, Psychology, and Human Resource. This was followed by respondents from Engineering (n = 65, 30.8%), which included Electrical, Electronic, Biomedical, and Mechanical Engineering. The remaining 28.0% (n = 59) consisted of participants from Science and Technology, such as Building Surveying, Architecture, Science, and Mathematics.

In Rasch measurement, sample adequacy is determined not only by the total number of respondents but also by the distribution of responses across rating scale categories. Linacre [32,33] suggested that at least 10 responses per rating scale category are recommended. Linacre [34] also notes that a sample of 80–100 well-targeted respondents provides stable item calibrations within approximately ±1 logit at a 95% confidence level. In the present study, all rating categories met this criterion, supporting the adequacy of N = 211 for generating stable Rasch estimates for descriptive purposes.

Nevertheless, although the sample size and response distribution were sufficient for Rasch analysis, the use of purposive sampling from a single university limits the generalisability of the findings. The results primarily reflect the assessment practices and GenAI-related behaviours of undergraduates within this specific institutional context and may not fully represent students from other universities or educational settings. Accordingly, the findings should be interpreted as context-specific and exploratory rather than universally generalisable.

Instrument

The instrument used in this study was a structured questionnaire developed to investigate students' parroting behaviour in AI-assisted academic writing, along with their general awareness of AI tools.

The questionnaire consisted of two sections (Table 2). Section A captured demographic information such as, awareness of generative AI, knowledge of access, and knowledge of use. These items were measured using a four-point Likert agreement scale (Strongly Agree = 4 to Strongly Disagree = 1). Section B measured parroting behaviour across three constructs: Intrinsic-Induced (10 items), Extraneous-Induced (10 items), and Ethically Rationalised (10 items). Responses were rated on a five-point Likert frequency scale ranging from Always (5) to Never (1).

Table 3 shows item reliabilities across all subscales were acceptable (>0.70), indicating consistent measurement of items within each construct. Item separation values also suggest adequate to strong differentiation among items (>2.0). For the Intrinsic-Induced construct, the item reliability was 0.98 with an item separation of 6.71, while the Extraneous-Induced construct recorded an item reliability of 0.98 and item separation of 6.66, both reflecting strong item stability. The Ethically Rationalised construct demonstrated an item reliability of 0.92 with an item separation of 3.42, indicating meaningful variation in item difficulty. Taken together, these indices meet recommended thresholds for Rasch measurement and support the suitability of the instrument for descriptive use in this study [35,36]

Table 1
Respondent distribution.

Field	N	%
Engineering	65	30.8%
Non-Engineering / Social Sciences	87	41.2%
Science and Technology	59	28.0%

Table 2
Items tabulation.

Section	Measure	No of Item	Likert-Type Scale
A	Demographic	3	4 Point Likert (Agreement)
	Awareness of Generative AI		Strongly Agree (4)
			Agree (3)
	Know How to Access		Disagree (2)
	Know How to Use		Strongly Disagree (1)
B	**Parroting Behaviour**		5 Point Likert Scale
	Intrinsic Load-Induced	10	(Frequency)
	Extraneous Load-Induced	10	Always (5)
	Ethically Rationalised	10	Often (4)
			Sometimes (3)
			Rarely (2)
			Never (1)

Table 3
Item and person reliability and separation for parroting behaviour subscales.

Parroting Behaviour	Item		Person	
	Reliability	Separation	Reliability	Separation
Intrinsic Induced	.98	6.71	.86	2.43
Extraneous Induced	.98	6.66	.85	2.41
Ethically Rationalised	.92	3.42	.80	2.01

Group behaviour across constructs

Group-based item functioning analysis was conducted to examine whether items performed equivalently across different respondent groups: Engineering, Non-Engineering/Social Sciences, and Science and Technology. This analysis evaluates the extent to which item difficulty estimates remain stable when comparing students from distinct academic fields, thereby identifying any potential measurement bias. The analysis ensures that observed differences in item responses reflect true variations in the underlying construct rather than systematic advantages or disadvantages for particular groups [37]

Table 4 summarises the T-scores, size indices, and standard errors for each construct across the three groups. For the Intrinsic-Induced construct, T-scores were 1112 (Engineering), 1857 (Non-engineering), and 1074 (Science and Technology); for Extraneous-Induced, the corresponding T-scores were 1260, 2043, and 1178; and for Ethically Rationalised, 1017, 1659, and 876. The size statistics index for all constructs and groups was consistently 0.00, and standard errors remained small (0.04–0.05) indicate no evidence of differential item functioning, meaning that the items behaved equivalently across academic fields. Such consistency is a strong indicator of the instrument's validity and fairness. It shows that the items are robust across disciplinary contexts and that the constructs are generalisable beyond a single group. For a behavioural measure that spans cognitive load responses and ethical rationalisation processes, this invariance is particularly important, as it confirms that patterns observed in the hierarchy are not discipline-specific but represent broader student tendencies.

Data collection

Data were collected online via a google form. The introductory section stated the purpose of the study, ensured anonymity, and reiterated participants' right to refuse or withdraw. No identifying information was collected.

Table 4
Group-based item functioning analysis.

PERSON class/group specification is: DIF=@FIELD ITEM class/group specification is: DPF=$S1W1

Construct	Engineering			Non-Engineering			Science and Technology		
	T. SCORE	SIZE	S.E	T. SCORE	SIZE	S.E	T. SCORE	SIZE	S.E
Intrinsic-Induced	1112	0.00	0.05	1857	0.00	0.04	1074	0.00	0.05
Extraneous-Induced	1260	0.00	0.05	2043	0.00	0.04	1178	0.00	0.05
Ethically Rationalised	1017	0.00	0.04	1659	0.00	0.04	876	0.00	0.05

Results

Awareness, access, and usage of AI-based writing tools

Table 5 presents the distribution of responses regarding participants' awareness and use of AI-based writing tools. A large majority of students reported high awareness of generative AI tools (e.g., ChatGPT) for academic writing, with 77.3% strongly agreeing and 19.0% agreeing. Only a small proportion (3.3%) disagreed, while 0.5% strongly disagreed. Similarly, most students indicated they know how to access AI-based writing tools, with 68.7% strongly agreeing and 26.1% agreeing. A small percentage (4.7%) disagreed and (0.5%) strongly disagreed, suggesting some gaps in accessibility knowledge. In terms of usage skills, 71.6% strongly agreed and 20.9% agreed that they know how to use AI-based writing tools. Only 5.2% disagreed, while 2.4% strongly disagreed.

Person–item mean distribution across constructs

The person–item distribution across constructs provides an overview of how respondents' endorsement levels align with the difficulty of items within each subscale. By comparing the mean person measures against the item mean set at 0.00 logits, this analysis illustrates how well each construct targets the sample and highlights relative tendencies in endorsing intrinsic-induced, extraneous-induced, and ethical-induced parroting behaviours.

Fig. 1 illustrates the distribution of person mean measures across the three constructs relative to the item mean (0.00 logits). The Extraneous-Induced Parroting construct was positioned slightly above the item mean (+0.18 logits), indicating that students, on average, displayed a comparatively higher tendency to rely on AI when encountering unclear or poorly structured assignments. In contrast, both Intrinsic-Induced Parroting (−0.14 logits) and Ethically Rationalised Parroting (−0.38 logits) fell below the item mean.

Item hierarchy (intrinsic load induced)

Fig. 2 presents the Rasch item map for intrinsic-induced parroting behaviours, displaying item difficulty along the logit scale. Using the mean item measure of 0.00 logits and the standard deviation of 0.58 logits as reference points, items located below –0.58 logits represent behaviours that are easier for students to endorse. Items falling within one standard deviation of the mean (–0.58 to +0.58 logits) reflect *moderately endorsed* behaviours, aligning with the overall difficulty level of the construct. Items positioned above +0.58 logits exceed one standard deviation from the mean and represent behaviours that are *harder to endorse*, requiring a stronger inclination toward intrinsic-induced parroting before respondents agree with them. This pattern illustrates how the distribution of item difficulties differentiates the relative ease or challenge of endorsing each behaviour within the construct.

Table 5
Agreement levels on awareness, access, and usage of AI-based writing tools.

Item	Agreement (%)			
	1 (Strongly Disagree)	2 (Disagree)	3 (Agree)	4 (Strongly Agree)
I am aware that AI tools such as (e.g. Chatgpt) can assist with academic writing.	1 (0.5%)	7 (3.3%)	40 (19.0%)	163 (77.3%)
I know how to access AI-based writing tools.	1 (0.5%)	10 (4.7%)	55 (26.1%)	145 (68.7%)
I know how to use AI-based writing tools.	5 (2.4%)	11 (5.2%)	44 (20.9%)	151 (71.6%)

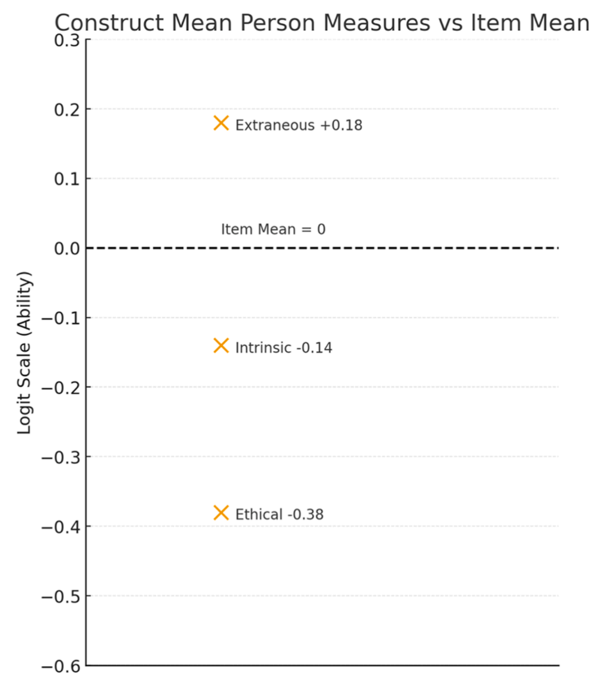

Fig. 1. Comparative mean person–item distribution.

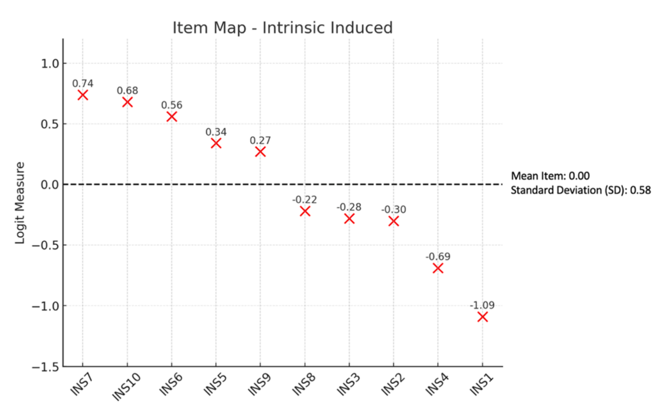

Fig. 2. Item hierarchy (intrinsic load-induced).

Easier endorsement

The items most easily endorsed include INS1 (“I use AI-generated text directly when the assignment topic feels too complex”, –1.09 logits) and INS4 (“I paraphrase AI output minimally when I cannot make sense of the full question”, –0.69 logits). Their positions well below the -0.58 indicate that, within the intrinsic-induced construct, students are comparatively more willing to admit relying on AI-generated text when they feel overwhelmed by task demands. In practical terms, these behaviours require less “activation” of intrinsic-induced parroting tendencies for respondents to agree with them, suggesting that using AI as a direct shortcut in challenging tasks is a relatively common strategy in this sample.

Moderate endorsement

Several items fall within this range. For instance, INS2 (“I insert AI text into my assignment when the task involves too many unfamiliar concepts,” –0.30 logits) and INS3 (“I rely on AI content without changes when I do not know how to begin writing,” –0.28 logits) indicate reliance on AI when students feel uncertain about how to initiate their work or structure their ideas.

Meanwhile item INS8 (“I use AI content that sounds professional to cover up my confusion with the topic,” –0.22 logits), item INS9 (“When the task is too hard, I submit AI-generated writing to make it look like I understand,” 0.27 logits) and INS5 (“I believe AI-generated content is better than mine, so I use it without making significant changes,” 0.34 logits) reflect a more sophisticated form of reliance, where students use AI not only to complete tasks but also to enhance the perceived quality or intellectual depth of their work.

Lastly, INS6 (“I skip outlining my ideas when I use AI to help with writing difficult topics,” 0.56 logits) illustrates how intrinsic cognitive strain can interrupt students’ ability to plan and structure their ideas. Collectively, these items reflect behaviours that students engage in situationally, often when facing cognitive uncertainty, feeling underprepared, or perceiving AI as a tool that compensates for gaps in their understanding or ability.

Less likely endorsement

Items positioned above one standard deviation from the mean (> +0.58 logits) represent behaviours that are less likely to be endorsed, reflecting actions that only students with a stronger tendency toward intrinsic-induced parroting admit to engaging in. Two items fall into this category. INS10 (“I use AI tools to make my assignments appear thoughtful even when I do not fully understand them,” 0.68 logits) captures a more deliberate form of reliance on AI, where students strategically enhance the perceived depth of their work despite limited comprehension. Similarly, INS7 (“I do not revise AI-generated responses because the assignment feels mentally exhausting,” 0.74 logits) reflects a high level of dependence driven by mental fatigue or cognitive overload, where students bypass revision entirely. Their positions above +0.58 logits indicate that these behaviours require a comparatively stronger inclination toward intrinsic-induced parroting before students are willing to endorse them. In other words, although students may frequently rely on AI for support in challenging tasks, only a smaller subset acknowledges engaging in these more intensive or effort-avoiding behaviours.

Item hierarchy (extraneous load induced)

Fig. 3 presents the Rasch item map for extraneous-induced parroting behaviours, displaying the relative difficulty of each item along the logit scale. Using the mean item measure of 0.00 logits and the standard deviation of 0.58 logits as reference points, items located below –0.58 logits represent behaviours that are easier for students to endorse, typically reflecting actions taken in response to external pressures such as limited time or heavy workload. Items falling within one standard deviation of the mean (–0.58 to +0.58 logits) reflect *moderately endorsed* behaviours, indicating tendencies that students acknowledge in certain situations but not as consistently as the easier items. Items positioned above +0.58 logits exceed one standard deviation from the mean and therefore represent *harder-to-endorse* behaviours, requiring a stronger inclination toward extraneous-induced parroting before respondents agree with them. This distribution demonstrates how the item hierarchy differentiates the relative ease or challenge of endorsing each behaviour within the extraneous-induced construct.

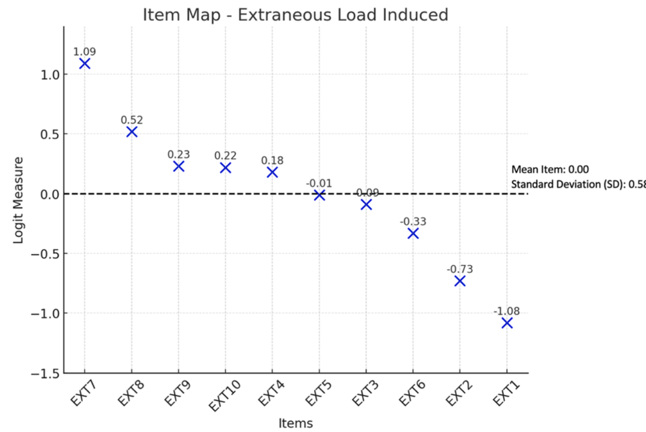

Fig. 3. Item hierarchy (extraneous load-induced).

Easier endorsement

Items falling more than one standard deviation below the mean (logits < –0.58) reflect behaviours that students most readily admitted when experiencing extraneous cognitive load. Two items; EXT1 (–1.08 logits) and EXT2 (–0.73 logits) were positioned in this category. EXT1 (“I use AI content in my work when the instructions are hard to follow.”) and EXT2 (“I use AI-generated paragraphs as-is when I cannot understand the assignment requirements”) indicate a strong tendency to rely on AI when assignment expectations are unclear or difficult to interpret. Their comparatively low logit values suggest that these behaviours represent common coping mechanisms, with students turning to AI to navigate ambiguity or incomprehensible instructions. These behaviours require minimal extraneous-induced prompting, making them the most easily endorsed within the construct.

Moderate endorsement

Items located within one standard deviation of the mean (–0.58 to +0.58 logits) represent behaviours that students acknowledge under certain conditions but do not endorse as consistently as the easier items. This group includes EXT6 (“When writing requirement is high, I use AI to generate content for assignment”, –0.33 logits), EXT3 (“I use AI content as if it were my own when I have too many assignments”, –0.09 logits), EXT5 (“When a deadline is approaching, I focus on changing just a few words from the AI output”, –0.01 logits), EXT4 (“I slightly reword AI-generated content and submit it when the assignment guidelines are unclear”, 0.18 logits), EXT10 (“If I run out of time, I rely on AI to produce work that seems well-structured and original. “, 0.22 logits), and EXT9 (“I use AI to complete writing quickly when I do not understand how to structure the assignment”, 0.23 logits). These behaviours capture situational reliance on AI driven by external pressures such as high workload, impending deadlines, vague requirements, or difficulty structuring assignments.

Meanwhile, EXT8 (“I avoid thinking critically when AI already provides a full answer under time pressure") sits at the upper boundary of this range, reflecting a behavioural shift toward cognitive offloading. Overall, items in this band illustrate behaviours that are neither frequent defaults nor rare occurrences, but rather responses enacted when external task demands intensify.

Less likely endorsement

Items positioned above one standard deviation from the mean (logits > +0.58) reflect behaviours that students are least willing to endorse. Only one item, EXT7 (1.09 logits), fell into this category. EXT7 ("When assignment instructions are vague, I allow AI to do most of the writing without my input") represents a more substantial level of reliance on AI, indicating near-total delegation of writing responsibilities under extraneous load. Its high logit value suggests that this behaviour requires the strongest extraneous-induced prompting and is endorsed only by a smaller subset of respondents. This pattern indicates that while students often rely on AI to navigate confusing or demanding tasks, fully surrendering authorship to AI without contributing their own ideas remains a comparatively uncommon practice.

Item hierarchy (ethical rationalised)

Fig. 4 illustrates the Rasch item–person map for ethical rationalisation–induced parroting behaviours, with item difficulty plotted along the logit scale. Using the mean item measure of 0.00 logits and the standard deviation of 0.29 logits as interpretive thresholds, items below –0.29 logits are classified as *easier to endorse*, items between –0.29 and +0.29 logits as *moderately endorsed*, and items exceeding +0.29 logits as *less likely to be endorsed*. This framework provides a clear differentiation of behavioural tendencies related to ethical rationalisation in the context of AI-assisted academic work.

Easier endorsement (negative logits)

Items falling below –0.29 logits represent ethical-induced behaviours that students are more inclined to endorse. Three items were located in this range: ETC8 (–0.32 logits), ETC4 (–0.38 logits), and ETC2 (–0.39 logits). These items reflect behaviours such as assuming AI use is acceptable when no explicit rules are provided ("If there are no clear rules about AI usage, I assume it is acceptable to use it freely," ETC8, –0.32 logits), withholding disclosure when institutional or course policies are unclear ("If there are no clear policies, I do not mention AI assistance in my work," ETC4, –0.38 logits), and avoiding citation when assignment guidelines do not address AI usage ("I avoid mentioning AI use when assignment instructions do not specify citation rules," ETC2, –0.39 logits). Their comparatively low logit values indicate that students readily rationalise nondisclosure of AI use when external guidance is uncertain, suggesting that ambiguity around rules and expectations strongly facilitates ethical rationalisation and non-transparent practices.

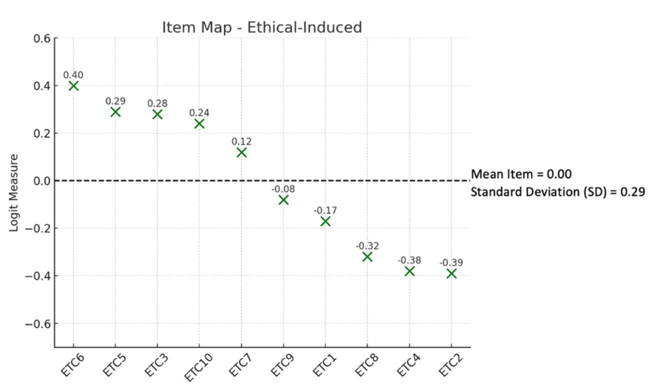

Fig. 4. Item hierarchy (ethical rationalised).

Moderate endorsement

Items positioned between –0.29 and +0.29 logits represent ethical-induced behaviours that students endorse under certain circumstances but not as consistently as the easier items. Six items were in this range: ETC1, ETC9, ETC7, ETC10, ETC3, and ETC5. These items reflect behaviours such as choosing not to cite AI-generated content when it appears generic or obvious ("I do not cite AI tools when the content seems generic or obvious," ETC1, –0.17 logits) and normalising AI use because peers do the same ("I believe it is fine to use AI to complete difficult tasks, especially when others do the same," ETC9, –0.08 logits).

Students also report feeling comfortable relying on AI due to perceived low detection risk ("I know it is hard to detect AI use, so I feel safe using it without making changes," ETC7, 0.12 logits) and submitting AI-generated work without acknowledgment ("I submit AI-generated work without acknowledgment, assuming that my lecturer will not notice," ETC10, 0.24 logits).

Additional behaviours in this category include uncertainty about proper citation ("I submit content from AI tools without citing them because I am unsure how to do it properly," ETC3, 0.28 logits) and a belief that AI-generated material does not constitute plagiarism ("I do not consider it plagiarism if I use content generated by AI tools in my assignments," ETC5, 0.29 logits). Together, these items illustrate ethically ambiguous practices that emerge when students weigh convenience, perceived norms, and uncertainty about proper academic procedures.

Harder endorsement

Only one item exceeded the +0.29-logit threshold, indicating a behaviour that students are least willing to endorse. ETC6 reflects a more explicit rationalisation of nondisclosure, capturing the belief that AI-generated content does not require citation because it lacks a human author ("I believe AI-generated content does not need to be cited because it has no human author," ETC6, 0.40). Its comparatively higher logit value indicates that students generally resist endorsing this stance, suggesting that outright dismissal of citation norms represents a more extreme form of ethical rationalisation within this construct.

Discussion

Extraneous-load induced parroting

The Rasch analysis highlights a clear pattern across the three dimensions of parroting behaviour. The extraneous dimension shows the highest average endorsement, meaning that students are most likely to use AI-generated text when they face external pressures such as tight deadlines, unclear instructions, or heavy workloads. This finding is consistent with cognitive load theory (CLT) which explains that poorly designed tasks can overload students' working memory and push them toward surface-level strategies [17,38,39]. In this context, AI-related parroting should not be viewed simply as disengagement or academic dishonesty; rather, it emerges as a compensatory strategy students adopt when instructional design unintentionally overwhelms their working memory [3,7,24]. When assignment clarity is low or deadlines converge, students may perceive AI tools not as shortcuts, but as necessary scaffolds to manage competing academic demands.

A closer look at the item difficulty hierarchy reveals a clear pattern in how students respond to these pressures. The most easily endorsed behaviours indicate that when assignment requirements are confusing or instructions lack specificity, students resort to AI as a compensatory strategy. This is consistent with CLT which emphasizes that ambiguous or poorly structured tasks create unnecessary processing demands unrelated to the actual learning goal [17,18,25]. When students report that they submit AI content “as-is” under such conditions, it underscores how extraneous load can redirect effort away from meaningful learning toward superficial task completion. Rather than investing effort in interpreting ambiguous expectations, students, consistent with Dual-Process Theory (DPT) bypass deeper cognitive processing (System 2) and instead rely on the automatic, heuristic responses (System 1) by outsourcing their work to AI tools [16,24]

A second set of behaviours reflects reliance on AI in response to situational pressures such as approaching deadlines, multiple competing assignments, demanding writing requirements, or difficulty structuring ideas. Students report using AI to accelerate writing, to navigate dense or vague guidelines, or to produce text that appears coherent when they cannot organise their thoughts under pressure. Unlike the more automatic reliance observed in the easily endorsed behaviours, these patterns emerge when external task demands escalate and students feel unable to allocate sufficient cognitive resources to meet expectations. This aligns with prior research showing that when extraneous load interacts with heavy workload, learners adopt “satisficing” strategies meeting minimum requirements with minimal effort [40,41]. Consequently, AI is used strategically to maintain productivity, reduce the time needed to generate or revise content, or meet performance standards with minimal cognitive investment. Here, the issue is not lack of capability, but rather misalignment between the demands of the task and the cognitive resources available to students.

A more concerning behavioural tendency appears when students describe relinquishing critical thinking to AI under time pressure or allowing AI to produce the majority of the writing when instructions are vague. These behaviours signal a transition from targeted assistance to broader cognitive offloading [42–44], where AI is used not only to support writing but also to replace central components of the writing process. Although fewer students endorse these behaviours, their presence indicates that extraneous load can create conditions where learners begin to withdraw from meaningful cognitive participation. Such reliance suggests that when instructional demands exceed students' perceived capacity for engagement, AI becomes a surrogate writer rather than a supplemental tool.

Taken collectively, the extraneous-induced construct highlights how avoidable instructional factors play a central role in prompting parroting behaviours [18,39]. When assignment guidelines are unclear, expectations are ambiguous, or workload feels unmanageable, students shift toward AI-driven shortcuts to manage the external pressures placed upon them. Importantly, reliance on AI in these situations is not necessarily a reflection of limited ability or motivation, but rather a response to the cognitive disruption caused by poorly structured or insufficiently supported tasks.

These findings underscore a key pedagogical implication: reducing extraneous cognitive load can substantially diminish the need for students to adopt AI-based coping strategies [17]. For instance, when lecturers provide explicit instructions, structured assignment scaffolds, and timely formative feedback [45,46], students experience fewer cognitive pressures that push them toward superficial shortcuts. By minimising unnecessary cognitive strain and clarifying the pathways to task completion, educators can foster deeper engagement and create learning environments that support authentic academic effort rather than reliance on parroting behaviours.

Intrinsic-load induced parroting

Academic writing involves complex cognitive skills such as synthesizing ideas and using disciplinary language. Intrinsic load–induced parroting arises when students perceive the task as exceeding their own ability. When topics feel abstract, concepts are unfamiliar, or the assignment requires synthesis they are not confident performing, students experience intrinsic cognitive overload. In these moments, they tend to believe that AI can produce better work than they could on their own [47,48], which leads them to bypass deeper reasoning and rely uncritically on AI-generated content as a way to cope with the demands of the task.

The item hierarchy results show that the most readily acknowledged behaviours occur at the initial stages of task engagement, where students face uncertainty about how to interpret questions or how to begin formulating ideas. In such moments, students often turn to AI-generated text as a way to reduce the cognitive friction associated with understanding complex topics. Rather than struggling through conceptual ambiguity, they use AI as an immediate scaffold, allowing them to bypass the mental effort required for early comprehension and idea generation [49–51]. This pattern aligns with cognitive load theory, which posits that learners experiencing high intrinsic load are more likely to seek external aids that alleviate the burden on working memory.

As writing tasks become more conceptually demanding, reliance on AI evolves into more nuanced forms. Students report using AI to fill gaps in understanding, articulate unfamiliar concepts, or supply language that feels more polished than what they believe they can produce independently. These behaviours are situationally driven, often triggered when students lack the confidence or capacity to engage fully with the assignment [49,52]. When students doubt their ability to synthesise disciplinary ideas or express them with sufficient clarity, AI becomes a compensatory mechanism that stabilises their performance. This is consistent with scholarship showing that low writing self-efficacy increases the use of external supports and shortcuts [51,53], particularly when students face tasks requiring specialised vocabulary or higher-order reasoning [54,55].

The least acknowledged behaviours involve more profound forms of cognitive disengagement. These include relying on AI to produce work that conveys understanding the student does not possess, or skipping revision entirely because the task feels mentally exhausting. These patterns signal a shift from targeted support to near-total cognitive offloading, where students relinquish ownership of the writing process [25,56]. Although these behaviours are less frequently endorsed, their presence is there: they illustrate how overwhelming intrinsic load can push some students toward avoidance rather than assistance [56,57], resulting in limited interaction with the cognitive processes that underlie learning.

Overall, the intrinsic-induced construct reveals a continuum of behaviours shaped by cognitive strain, writing insecurity, and perceived deficits in conceptual understanding. Students tend to rely on AI when they cannot interpret a task or when the complexity of disciplinary content overwhelms their working memory, and this reliance becomes more sophisticated as cognitive challenges deepen. For some, AI serves as pragmatic support to initiate writing; for others, it becomes a tool to enhance the appearance of competence or, in a smaller subset, a means of withdrawing from cognitive effort altogether. This pattern is consistent with earlier work showing that writing insecurity is a major driver of plagiarism-like behaviours [6,7,51], highlighting that academic integrity policies alone are insufficient. Students who lack confidence in their writing or conceptual abilities are more likely to view parroting as their only viable option. These findings underscore the need for targeted scaffolding, and instructional interventions that build students' confidence in engaging with complex ideas [51,52]. Strengthening these capacities may reduce dependence on AI-driven shortcuts and promote more sustained and meaningful cognitive participation in academic writing.

Ethical rationalised parroting

Ethical rationalisation–induced parroting refers to behaviours in which students justify the nondisclosure or improper use of AI by appealing to ambiguity, perceived norms, or personal interpretations of academic rules [58]. Rather than being driven by cognitive overload, these behaviours arise when students reinterpret ethical boundaries in ways that make AI reliance seem acceptable, excusable, or harmless [4, 15]. This construct captures the psychological mechanisms through which students normalise nondisclosure and minimise perceived wrongdoing in the context of generative AI.

Students most readily acknowledge behaviours that involve using ambiguity as a justification for nondisclosure. When institutional policies do not explicitly address AI use, or when assignment guidelines fail to clarify citation expectations, students report feeling comfortable omitting acknowledgment of AI-generated content. This pattern suggests that uncertainty around rules creates a permissive space where students interpret silence as implicit approval. Ethical boundaries become negotiable when learners believe responsibility lies not in adhering to academic norms but in the institution's ability to articulate them. Such reasoning reflects classic moral disengagement processes, in which individuals shift accountability away from themselves by claiming ambiguous standards or external omissions [27,42].

This finding raises an important concern: if students believe using AI without proper engagement is "normal," the issue is less about individual dishonesty and more about how academic norms are communicated [49]. Research suggests that students are more likely to act ethically when institutions clearly teach why integrity matters [3,13] and connect it to learning and personal growth [59,60]. Without helping students reflect on the ethical side of authorship, any rules or deterrents risk being ignored or followed only superficially.

The item hierarchy finding illustrates a clear gradient in how students rationalise the undisclosed use of AI tools, revealing important perspectives into the ethical dimension of emerging academic practices. At the easier end of the continuum, students' willingness to bypass disclosure is closely tied to contextual ambiguity. When institutional guidelines are vague or non-existent, or when detection appears unlikely, students perceive little ethical risk in using AI content without attribution [13,61]. Such patterns indicate that policy opacity and weak enforcement function as critical enabling conditions for non-disclosure.

A second set of behaviours highlights more situational forms of ethical rationalisation. Students describe choosing not to cite AI when its output appears generic, when they feel unsure about proper citation procedures, or when they perceive AI use as a common practice among peers. They also report a willingness to rely on AI because detection feels unlikely, which diminishes the perceived consequences of nondisclosure. These behaviours indicate that ethical decision-making is influenced by convenience, social norms, and perceptions of risk rather than by principled adherence to integrity [13,62]. In these cases, students do not reject academic values outright; instead, they interpret them flexibly in ways that reduce personal effort, limit vulnerability to punishment, or align with what they perceive others are doing.

In contrast, students are less willing to explicitly deny established plagiarism norms or challenge the principle that AI-generated content requires acknowledgement. The most difficult rationalisations involve outright claims that AI-assisted writing is not plagiarism or that it does not require citation because it lacks a human author. The relative reluctance to endorse such views highlights that most students still recognise a baseline of academic integrity [15,62,63] even when those contributions come from non-human sources. The boundary between acceptable and unacceptable behaviour is therefore not entirely eroded; instead, it becomes selectively expanded in areas where students feel uncertain, unsupported, or shielded from accountability.

In summary, the ethical rationalisation construct reveals how students navigate the moral grey areas surrounding generative AI. Their justifications emerge not from malicious intent but from ambiguity, perceived norms, and uncertainty about proper academic conduct [10,61]. This highlights a crucial implication: strengthening academic integrity in the age of AI requires more than punitive policies [1,3,61]. Clear guidance on citation, explicit communication about acceptable and unacceptable uses of AI, and transparent expectations for academic honesty are essential to limiting the space in which students morally rationalise nondisclosure. When institutions articulate expectations clearly and model ethical use, students are less likely to reinterpret boundaries in self-serving ways and more likely to engage with AI in ways that align with academic values.

Implications

The findings of this study offer important implications for theory, pedagogy, and institutional policy in higher education. The Rasch-derived hierarchies show that parroting behaviours emerge through a structured progression shaped by cognitive load, intuitive reasoning shortcuts, and ethical rationalisation. This progression demonstrates that inappropriate reliance on AI is rarely the result of intentional misconduct alone; instead, it reflects the interaction of intrinsic cognitive strain, environmental pressures, and students' efforts to morally justify their actions. Understanding these layered influences is crucial for designing interventions that meaningfully support students rather than merely penalising them.

From a theoretical perspective, the results highlight the need to understand AI misuse as a cognitive and motivational phenomenon rather than a purely moral one. Intrinsic load explains why students rely on AI when they feel unable to understand or synthesise complex ideas. Extraneous load accounts for patterns that emerge when assignment expectations are unclear or when the volume of academic work overwhelms students' ability to think through tasks carefully. Dual mode reasoning helps explain why students resort to rapid, low-effort responses when under strain, allowing AI to fill the space where reflective thought would ordinarily occur. Ethical rationalisation explains how students maintain a sense of personal integrity while engaging in questionable practices by reframing rules, shifting responsibility, or interpreting ambiguity as permission. Together, these perspectives suggest that AI misuse is not a single behaviour but a predictable progression shaped by context, cognition, and self-perception.

There are also important implications for academic integrity practices. When expectations surrounding AI are vague or inconsistent, students fill the gaps with their own interpretations, often in ways that justify nondisclosure. Policies should therefore be explicit, accessible, and aligned across courses and programmes. Ethical guidelines must be communicated clearly enough to eliminate uncertainty that students might use as justification for misleading practices. However, clarity alone is insufficient. Institutions must cultivate a culture in which responsible use of AI is consistently modelled and in which discussions of ethical reasoning accompany the development of academic skills.

At the policy level, the findings argue for a shift from reactive enforcement to proactive capacity building. Approaches that rely solely on detection or punishment overlook the cognitive and contextual pressures that lead students to rely on AI in the first place. Institutions should focus on creating learning environments that minimise avoidable sources of cognitive strain, strengthen students' sense of academic competence, and promote ethical self-awareness. Policies that position integrity as a shared responsibility rather than an individual obligation are more likely to reduce the conditions that give rise to rationalisation and misuse.

Overall, the implications of this work point toward the need for coordinated changes in teaching, institutional communication, and student support. Addressing intrinsic challenges, reducing extraneous pressures, guiding ethical reasoning, and clarifying expectations can collectively reduce the appeal of AI-driven shortcuts and foster more sustained and authentic learning.

Recommendations

While the present study offers descriptive information into students' intrinsic, extraneous, and ethical rationalisations for AI misuse, several limitations suggest important directions for future work. These recommendations aim to strengthen both the theoretical and practical implications of subsequent studies.

Firstly, this research was conducted within a single institution, which limits the generalisability of the Rasch-based findings on differential functioning and item hierarchy. Although the analysis demonstrated stable item behaviour across academic fields within the university, it remains unclear whether the same invariance would hold in institutions with different curricular structures, assessment cultures, or student profiles. Future studies should therefore replicate the Rasch differential analysis and item hierarchy across multiple institutions to determine whether the progression of intrinsic, extraneous, and ethical rationalisation behaviours reflects a broader pattern or is influenced by local academic contexts. Such cross-institutional validation would strengthen confidence in the construct stability of the instrument and provide a more comprehensive understanding of how students in diverse environments engage with generative AI.

Secondly, the current study relied exclusively on questionnaire data, which, although appropriate for establishing item hierarchy through Rasch analysis, limits the depth with which students' underlying reasoning can be understood. The purpose of the instrument was to measure behavioural tendencies and their structured progression across intrinsic, extraneous, and ethical dimensions, but the quantitative approach cannot fully capture the nuances of how students interpret task difficulty, negotiate cognitive pressures, or construct ethical justifications in real time. Future research would benefit from incorporating qualitative methods such as interviews, think-aloud protocols, or analysis of student writing processes to contextualise the hierarchical patterns identified in the Rasch model. Such methodological triangulation would provide richer insight into why certain behaviours are more foundational, how students transition from one behavioural level to another, and how cognitive and ethical factors interact during actual engagement with academic tasks.

Thirdly, the study focused solely on student behaviours and did not capture the perspectives of lecturers or the institutional policies that shape students' interpretations of acceptable AI use. In practice, students' ethical rationalisations and responses to cognitive load are strongly influenced by how instructors articulate expectations, structure assignments, and model responsible use of AI. Likewise, institutional policies often vary in clarity, enforcement, and alignment across departments, creating inconsistencies that students may use to justify nondisclosure. Future research should therefore examine lecturers' beliefs, teaching practices, and assessment designs, as well as the institutional policy landscape that frames AI usage. Understanding how these structural and pedagogical factors interact with student behaviours would allow researchers to identify systemic contributors to AI misuse and design interventions that operate at classroom and institutional levels rather than at the level of student behaviour alone.

These recommendations suggest a research agenda that moves beyond descriptive measurement toward deeper, contextually grounded understanding of AI-related academic behaviours. Future studies should expand across institutions to test the stability of the item hierarchy, combine quantitative measurement with qualitative inquiry to illuminate the cognitive and ethical mechanisms underlying each behavioural level, and include lecturer perspectives to capture the broader pedagogical ecosystem in which students make decisions about AI use. Institutional policy research should also be prioritised to identify how clarity, consistency, and cultural norms influence students' ethical interpretations. By integrating these directions, future scholarship can develop a more comprehensive model of AI engagement in higher education, one that accounts for cognitive constraints, instructional design, social norms, and institutional contexts. This integrated approach will support the development of interventions that not only discourage misuse but also strengthen authentic learning, ethical reasoning, and student agency in an AI-mediated academic environment.

Funding

This work was supported by the Universiti Teknologi Malaysia under the Potential Academic Staff Research University Grant (Q.J130000.2753.03K65).

Declaration

During the preparation of this manuscript, the author(s) used ChatGPT to assist with language refinement and grammar checking. After using this tool, the author(s) carefully reviewed and edited the content as necessary and assume full responsibility for the final version of the publication.

CRediT authorship contribution statement

Ibnatul Jalilah Yusof: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Zakiah Mohamad Ashari:** Validation, Conceptualization. **Lukman Hakim Ismail:** Resources, Project administration, Funding acquisition. **Mira Panadi:** Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M.R. King, A conversation on artificial intelligence, chatbots, and plagiarism in higher education, Cell Mol. Bioeng. 16 (1) (2023) 1–2, https://doi.org/10.1007/s12195-022-00754-8.
[2] C. Wang, S.J. Aguilar, J.S. Bankard, E. Bui, B. Nye, Writing with AI: what college students learned from utilizing ChatGPT for a writing assignment, Educ. Sci. 14 (9) (2024) 976, https://doi.org/10.3390/educsci14090976 (Basel).
[3] J. Crawford, M. Cowling, K.A. Allen, Leadership is needed for ethical ChatGPT: character, assessment, and learning using artificial intelligence (AI), J. Univ. Teach. Learn. Pract. 20 (3) (2023) 02, https://doi.org/10.53761/1.20.3.02.
[4] L.A. Gammoh, ChatGPT in academia: Exploring university students' risks, misuses, and challenges in Jordan, J. Furth. High. Educ. 48 (6) (2024) 608–624, https://doi.org/10.1080/0309877X.2024.2378298.
[5] Y.Y. Lo, D. Fung, X. Qiu, Assessing content knowledge through L2: mediating role of language of testing on students' performance, J. Multiling. Multicult. Dev. 44 (10) (2023) 1013–1028, https://doi.org/10.1080/01434632.2020.1854274.
[6] B.T. Gallant, M. Picciotto, G. Bozinovic, E. Tour, Plagiarism or not? Investigation of Turnitin®-detected similarity hits in biology laboratory reports, Biochem. Mol. Biol. Educ. 47 (4) (2019) 370–379, https://doi.org/10.1002/bmb.21236.
[7] Glatt, B. S. (1987). The cognitive consequences of parroting. (T-30189) [Doctoral dissertation, The University of Chicago]. ProQuest Dissertations & Theses Global.
[8] C.B. Paulston, M.N. Bruder, Teaching English as a second language, Tech. Proced. (1976). Retrieved from, https://files.eric.ed.gov/fulltext/ED153499.pdf.
[9] J. Tepperman, T. Stanley, K. Hacioglu, B. Pellom, Testing suprasegmental English through parroting, in: Proceedings of the Speech Prosody, Chicago, IL, USA, 2010, pp. 11–14. https://www.isca-archive.org/speechprosody_2010/tepperman10_speechprosody.pdf.
[10] A.B. Cyphert, Generative AI, plagiarism, and copyright infringement in legal documents, Minn. JL Sci. Tech. 25 (49) (2023). https://scholarship.law.umn.edu/mjlst/vol25/iss2/9.
[11] L. Stearns, Copy wrong: Plagiarism, process, property, and the law, Calif. Law Rev. 80 (1992) 513. https://www.jstor.org/stable/pdf/3480772.pdf.
[12] B.L. Frye, Should using an AI text generator to produce academic writing be plagiarism? Fordham Intell. Prop. Media Ent. LJ 33 (2022) 946. https://ssrn.com/abstract=4292283.
[13] A. Kleebayoon, V. Wiwanitkit, Artificial intelligence, chatbots, plagiarism and basic honesty: comment, Cell Mol. Bioeng. 16 (2) (2023) 173–174, https://doi.org/10.1007/s12195-023-00759-x.
[14] COPE Council (2023). COPE position - Authorship and AI - English. 10.24318/cCVRZBms.
[15] A.M. Jarrah, Y. Wardat, P. Fidalgo, Using ChatGPT in academic writing is (not) a form of plagiarism: what does the literature say, Online J. Commun. Media Technol. 13 (4) (2023) e202346, https://doi.org/10.30935/ojcmt/13572.
[16] D. Kahneman. Thinking, Fast and Slow, Farrar, Straus and Giroux, 2011. https://psycnet.apa.org/record/2011-26535-000.
[17] O. Chen, F. Paas, J. Sweller, A cognitive load theory approach to defining and measuring task complexity through element interactivity, Educ. Psychol. Rev. 35 (2) (2023) 63, https://doi.org/10.1007/s10648-023-09782-w.

[18] J. Sweller, Cognitive load theory, learning difficulty, and instructional design, Learn. Instr. 4 (4) (1994) 295–312, https://doi.org/10.1016/0959-4752(94)90003-5.
[19] S. Gupta, A. Jaiswal, A. Paramasivam, J. Kotecha, Academic writing challenges and supports: perspectives of international doctoral students and their supervisors, in: Frontiers in Education, 7, Frontiers Media SA, 2022 891534, https://doi.org/10.3389/feduc.2022.891534.
[20] L.P. Patac, A.V Patac Jr, Using ChatGPT for academic support: managing cognitive load and enhancing learning efficiency–A phenomenological approach, Soc. Sci. Humanit. Open 11 (2025) 101301, https://doi.org/10.1016/j.ssaho.2025.101301.
[21] M.F. Teng, M. Yue, Metacognitive writing strategies, critical thinking skills, and academic writing performance: a structural equation modeling approach, Metacogn. Learn. 18 (1) (2023) 237–260, https://doi.org/10.1007/s11409-022-09328-5.
[22] S. Mat Yusoff, A. Mohamad Marzaini, L. Hao, et al., Understanding the role of AI in Malaysian higher education curricula: an analysis of student perceptions, Discov. Comput. 28 (2025) 62, https://doi.org/10.1007/s10791-025-09567-5.
[23] I.M. Pepperberg, Cognition and communication in an African grey parrot (Psittacus erithacus): studies on a Nonhuman Nonprimate, nonmammalian subject. Language and Communication, Psychology Press, 2013, pp. 221–248.
[24] B.O. Omer, Why do students still plagiarize? Perceptions of EFL and non-EFL students on plagiarism, J. Univ. Hum. Dev. 8 (2) (2022) 54–60, https://doi.org/10.21928/juhd.v8n2y2022.pp54-60.
[25] L. Yan, V. Pammer-Schindler, C. Mills, A. Nguyen, D. Gašević, Beyond efficiency: empirical insights on generative AI's impact on cognition, metacognition and epistemic agency in learning, Br. J. Educ. Technol. 56 (2025) 1675–1685, https://doi.org/10.1111/bjet.70000.
[26] A. Bandura, Social cognitive theory of self-regulation, Organ. Behav. Hum. Decis. Process 50 (2) (1991) 248–287, https://doi.org/10.1016/0749-5978(91)90022-L.
[27] A. Abbas, A. Fatima, A. Arrona-Palacios, et al., Research ethics dilemma in higher education: impact of internet access, ethical controls, and teaching factors on student plagiarism, Educ. Inf. Technol. 26 (2021) 6109–6121, https://doi.org/10.1007/s10639-021-10595-z (Dordr).
[28] R.A. Vasquez, The ethical decision-making gap in student ethics: examining how university students approach ethical dilemmas, Int. J. Ethics Educ. 7 (1) (2022) 53–71, https://doi.org/10.1007/s40889-021-00133-3.
[29] D. Ayton, C. Hillman, K. Hatzikiriakidis, T. Tsindos, S. Sadasivan, S. Maloney, D. Illic, Why do students plagiarise? Informing higher education teaching and learning policy and practice, Stud. High. Educ. 47 (9) (2021) 1921–1934, https://doi.org/10.1080/03075079.2021.1985103.
[30] J.Á. De Lima, Á. Sousa, A. Medeiros, B. Misturada, C. Novo, Understanding undergraduate plagiarism in the context of students' academic experience, J. Acad. Ethics 20 (2) (2022) 147–168, https://doi.org/10.1007/s10805-021-09396-3.
[31] B. Mujtaba, Clarifying ethical dilemmas in sharpening students' artificial intelligence proficiency: dispelling myths about using AI tools in higher education, Bus. Ethics Leadersh. 8 (2) (2024) 107–127, https://doi.org/10.61093/bel.8(2).107-127.2024.
[32] J.M. Linacre, Understanding Rasch measurement: estimation methods for Rasch measures, J. Outcome Meas. 3 (4) (1999) 382–405. https://www.researchgate.net/profile/Alfred-Stenner/publication/12729624_Mapping_variables/links/577c092908ae355e74f169aa/Mapping-variables.pdf#page=92.
[33] J.M. Linacre, Optimizing rating scale category effectiveness, J. Appl. Meas. 3 (1) (2002) 85–106. https://www.researchgate.net/profile/John-Linacre/publication/11372384_Understanding_Rasch_measurement_Optimizing_rating_scale_category_effectiveness/links/65548bccb1398a779d8f59be/Understanding-Rasch-measurement-Optimizing-rating-scale-category-effectiveness.pdf.
[34] J.M. Linacre, Sample size and item calibration (or person measure) stability, Rasch Meas. Trans. 7 (4) (1994) 328. Retrieved from, http://www.rasch.org/rmt/rmt74m.htm.
[35] S. Dabaghi, F. Esmaielzadeh, C. Rohani, Application of Rasch analysis for development and psychometric properties of adolescents' quality of life instruments: a systematic review, Adolesc. Health Med. Ther. 11 (2020) 173–197, https://doi.org/10.2147/AHMT.S265413.
[36] A. Tennant, P.G. Conaghan, The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum. 57 (2007) 1358–1362, https://doi.org/10.1002/art.23108.
[37] Jr.G. Engelhard, Using item response theory and model—data fit to conceptualize differential item and person functioning for students with disabilities, Educ. Psychol. Meas. 69 (4) (2009) 585–602, https://doi.org/10.1177/0013164408323240.
[38] R. Mulenga, H. Shilongo, Academic integrity in higher education: understanding and addressing plagiarism, Acta Pedagogia Asiana 3 (1) (2024) 30–43, https://doi.org/10.53623/apga.v3i1.337.
[39] A. Skulmowski, K.M. Xu, Understanding cognitive load in digital and online learning: a new perspective on extraneous cognitive load, Educ. Psychol. Rev. 34 (1) (2022) 171–196, https://doi.org/10.1007/s10648-021-09624-7.
[40] D.R. Blazek, J.T. Siegel, Preventing satisficing: a narrative review, Int. J. Soc. Res. Methodol. 27 (6) (2024) 635–648, https://doi.org/10.1080/13645579.2023.2239086.
[41] T.C. Cheng, Y. Zhang, Y.H. Chou, V. Koshy, T.W. Li, K. Karahalios, H. Sundaram, Organize, then vote: Exploring cognitive load in quadratic survey interfaces, in: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, 2025, pp. 1–35, https://doi.org/10.1145/3706598.3714193.
[42] P. Ayala-Enríquez, J. Guerrero-Dib, Moral disengagement leading to social acceptance of academic misconduct: a predictor of behavior. Handbook of Academic Integrity, Springer, 2023, pp. 1–24.
[43] C.K.Y. Chan, Understanding AI guilt: The development, pilot-testing, and validation of an instrument for students, Educ. Inf. Technol. (2025) 1–20, https://doi.org/10.1007/s10639-025-13629-y (Dordr).
[44] A.H. Perry, D.A. Rettinger, J.M. Stephens, E.M. Anderman, M.L. McTernan, H. Tatum, T. Bertram Gallant, From institutional climate to moral attitudes: examining theoretical models of academic misconduct, Ethics Behav. (2025) 1–18, https://doi.org/10.1080/10508422.2025.2514577.
[45] D.J. Nicol, D. Macfarlane-Dick, Formative assessment and self-regulated learning: a model and seven principles of good feedback practice, Stud. High. Educ. 31 (2) (2006) 199–218, https://doi.org/10.1080/03075070600572090.
[46] S. Quinn, Using "how to…" videos in feedforward practices to support the development of academic writing, J. Empower. Teach. Excell. 5 (3) (2022) 42–54. https://digitalcommons.usu.edu/jete/vol5/iss3/6.
[47] A. Morrison, Meta-writing: AI and writing, Compos. Stud. 51 (1) (2023). https://rws511.pbworks.com/w/file/fetch/154174671/Morrison%20Meta%20Writing.pdf.
[48] W. Simon, Distinguishing between student and AI-generated writing: a critical reflection for teachers, Metaphor (3) (2023) 16–23. https://search.informit.org/doi/10.3316/informit.274843605275892.
[49] R.W. Black, B. Tomlinson, University students describe how they adopt AI for writing and research in a general education course, Sci. Rep. 15 (1) (2025) 8799, https://doi.org/10.1038/s41598-025-92937-2.
[50] S. Koudsia, M. Kirchner, Reducing cognitive overload for students in higher education: a course design case study, J. High. Educ. Theory Pract. (10) (2024) 24, https://doi.org/10.33423/jhetp.v24i10.7382.
[51] J. Rodríguez-Ruiz, I. Marín-López, R. Espejo-Siles, Is artificial intelligence use related to self-control, self-esteem and self-efficacy among university students? Educ. Inf. Technol. 30 (2) (2025) 2507–2524, https://doi.org/10.1007/s10639-024-12906-6 (Dordr).
[52] A.S. Nelson, P.V. Santamaría, J.S. Javens, Students' perceptions of generative AI use in academic writing, in: Proceedings of the Conference Proceedings. Innovation in Language Learning 2024, 2024, https://doi.org/10.3390/educsci15050611.
[53] E. Shmeleva, T. Semenova, Academic dishonesty among college students: academic motivation vs. contextual factors, Вопросы образования 3 (2019) 101–129. https://cyberleninka.ru/article/n/academic-dishonesty-among-college-students-academic-motivation-vs-contextual-factors/viewer.
[54] E. Bensalem, R. Harizi, A. Boujlida, Exploring undergraduate students' usage and perceptions of AI writing tools, Glob. J. Foreign Lang. Teach. 14 (2) (2024) 53–65, https://doi.org/10.18844/gjflt.v14i2.9344.
[55] S. Bongsu, W.N.M Nik Mohamed, W.N Wan Mohamed, Determinants of students' intention to use AI-powered writing tools in academic writing, Int. J. Res. Innov. Soc. Sci. (IJRISS) 9 (03) (2025) 5616–5630, https://doi.org/10.47772/IJRISS.2025.903SEDU0411.
[56] Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.H., Beresnitzky, A. V., Maes, P. (2025). Your brain on ChatGPT: accumulation of cognitive debt when using an ai assistant for essay writing task. arXivLabs 10.48550/arXiv.2506.08872.
[57] K. Blackwell-Starnes, I prefer my own writing": Engaging first-year writers' agency with generative AI, Thresholds Educ. 48 (1) (2025) 25–39. https://files.eric.ed.gov/fulltext/EJ1468038.pdf.
[58] P.V. Sysoyev, Ethics and AI-plagiarism in an academic environment: students' understanding of compliance with author's ethics and the problem of plagiarism in the process of interaction with generative artificial intelligence, Vyss. Obraz. v Ross. Higher Educ. Russ. 33 (2) (2024) 31–53, https://doi.org/10.31992/0869-3617-2024-33-2-31-53.
[59] S. Atlas, ChatGPT for Higher Education And Professional Development: A Guide To Conversational AI, University of Rhode Island, 2023. Retrieved from, https://digitalcommons.uri.edu/cba_facpubs/548.
[60] S. Grassini, Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings, Educ. Sci. 13 (7) (2023) 692, https://doi.org/10.3390/educsci13070692 (Basel).
[61] J.M. Higgs, A. Stornaiuolo, Being human in the age of generative AI: young people's ethical concerns about writing and living with machines, Read. Res. Q. 59 (4) (2024) 632–650, https://doi.org/10.1002/rrq.552.
[62] Graf, A. E. (2024). "You Do You": Digital-Native College Students' Perceptions of Cheating and Plagiarism (Order No. 30812633). Available from ProQuest Dissertations & Theses Global. (2895144273). https://vpn.utm.my/dissertations-theses/you-do-digital-native-college-students/docview/2895144273/se-2.
[63] P. Shukla, S. Naik, I. Obi, J. Backus, N. Rasche, P. Parsons, Rethinking citation of AI sources in student–AI collaboration within HCI design education, in: Proceedings of the 7th Annual Symposium on HCI Education, 2025, pp. 1–8, https://doi.org/10.1145/3742901.3742909.