From Creation Myths to Technical AI Agency: Using Engineering Lenses to Explore Productive Failure of Unguided Generative AI Use in Medical Students

Andrea Carolina Velez-Parra, Universidad Antonio Nariño

Andrea Vélez is a Medical Doctor with a Master's degree in Human Anatomy, seven years of experience as an Assistant professor and six as Academic Coordinator in the College of Medicine from Universidad Antonio Nariño, Colombia, and an increasing interest in innovative forms of teaching basic science to first- and second-year medical undergraduates. Recently, interested in using AI in medical teaching, utilizing engineering education frameworks.

Anyerson Cuervo-Basurto, Virginia Tech Department of Engineering Education

Anyerson Cuervo is a Ph.D. Engineering Education student at Virginia Tech. Anyerson has a background in mechanical engineering and a master's degree in materials science. He has work experience in engineering roles at Colombian companies and has taught materials science and technical drawing courses at Colombian universities for four years. His current research interests include teacher identity in engineering faculty, cultural and institutional challenges in engineering teaching, and improvement of faculty development programs.

From Creation Myths to Technical AI Agency: Using Engineering Lenses to Explore Productive Failure of Unguided Generative AI Use in Medical Students

Abstract

Generative Artificial Intelligence (GAI) is reshaping professional education, yet it poses significant challenges for academic implementation. These include fear of academic dishonesty, misconceptions about its functioning, and ethical risks arising from uncritical reliance on automated outputs. In Engineering Education, there has been growing recognition of GAI as a co-pilot, with its use through “prompt engineering”, a process analogous to the engineering design cycle. In contrast, in other STEM areas, such as foundational Medical Education, where ethical and critical thinking are also fundamental learning objectives, traditional pedagogical paradigms often lack the analytical frameworks to evaluate these pedagogical interventions. In addition, given the increasing use of GAI in STEM, it remains unclear whether the mechanisms through which students-GAI interactions develop technical agency and ethical judgment, or whether these can spontaneously emerge in non-technical cohorts through unguided “productive failure”. This study addresses this gap by exploring how unguided interaction with GAI triggers iterative design behaviors and epistemic validation strategies in STEM medical students. Using the Productive Failure Engineering Education framework as a theoretical lens, we examined the procedural experiences of five medical students (semesters I–VI) tasked with generating visual conceptual representations using GAI without prior technical training. Using an exploratory qualitative thematic analysis, we identified how the initial lack of guidance forced students to confront the algorithm’s hidden nature. Participants initially treated GAI as an authoritative source, which led to failure and annoyance, inducing them to try again. This Productive Failure catalyzed a shift in cognitive strategy: students spontaneously adopted iterative debugging cycles, refined constraints (prompt specificity), and developed verification protocols to check GAI outputs against domain truth. The results suggest that unguided GAI interaction fosters a “designer mindset” in medical students, mirroring iterative behaviors typical of engineering undergraduates. Our study contributes to Engineering Education by demonstrating how multidisciplinary frameworks can support AI literacy and critical thinking across professional fields. Lastly, we provide insight into how engineering constructs such as Productive Failure and Iteration operate outside formal engineering curricula, revealing which elements of technical agency in STEM undergraduate students emerge spontaneously.

Keywords: Generative Artificial Intelligence, Medical Education, Iteration Theory, Productive Failure, Design Theory, Ethical Reflection, Multidisciplinary Engineering.

Introduction

The rapid emergence of Generative Artificial Intelligence (GAI) has triggered a paradigm shift in STEM Higher Education. In Medical Education, GAI offers transformative potential, ranging from the generation of synthetic clinical cases to the automation of complex documentation [1]. However, its integration into the biomedical classroom is often met with institutional resistance. As Medical Education in the Global South remains deeply rooted in traditional instructional methods, often described by students as “old-fashioned” [2]. The prevailing concern in medical faculty, which mirrors many concerns in other STEM areas, centers on academic integrity and the risk of “cognitive erosion,” in which students might rely on AI to provide immediate answers without clinical reasoning or critical thinking [3].

While medical students are increasingly using GAI in their daily academic lives, there is a delay regarding faculty efforts to integrate these tools into the classroom or train students in their critical use. This led to a significant gap: the availability of advanced technological tools, but the requirement of adequate pedagogical frameworks to use them critically [3]. In the absence of an official “medical education guide for AI”, most students are left to operate in a technical vacuum. Research on AI in medical education focuses on “Prompt Engineering” as a technical skill to be taught via direct instruction [1], [2]. Current Prompt Engineering has been taught as a series of critical steps that must be followed and planned accordingly, a structured and, in some sense, rigid framework for non-engineering students [2]

Although Prompt Engineering is critical for achieving desired outcomes, another way to address this technical skill is to interpret the value of GAI in professional formation as lying not in mastering technical instructions but in the reflective process triggered by the tool’s failures and complexities. This reflective process of students’ interactions mirrors principles of Productive Failure [4] and the Iteration Theory of Engineering Education, where effective prompting requires defining constraints, testing outputs, and refining inputs based on failure analysis [5]. It is possible that, in the same way, when non-engineers, such as medical students, are forced to navigate the complexity of new technologies, such as GAI, they mimic these iteration steps. When they navigate it without a map, they would engage in a high-stakes cognitive struggle that might naturally elicit design thinking [6] and transform them from passive consumers to active verifiers.

For those reasons, we aim to test this hypothesis and create a unique opportunity to observe if engineering logic, such as iterative refinement and validation, spontaneously emerges in non-technical disciplines. Thus, we designed an activity for medical students in which they used GAI without prior technical training, focusing on their experiences with early failure, the recursive need for iteration, and the challenge of confronting the algorithm’s hidden nature.

By framing these experiences through the lens of Engineering Education frameworks, our study addresses the following research question: How do medical students reflect on their iterative refinement and technical agency through the “Productive Failure” in unguided GAI interaction?

Background

The “No-Error” Paradigm in Medical Education

Medical education has historically been governed by a culture of clinical perfection, in which the primary objective is to minimize errors to ensure patient safety [7]. This “no-error” ethos, while essential for clinical practice, creates a rigid pedagogical environment where trial-and-error is often penalized rather than utilized as a learning mechanism [3], [8]. Consequently, medical students are trained to seek the “correct” answer on the first attempt, leaving little room for iterative experimentation [9], [10].

One foundational pedagogical strategy designed for medical students to practice “failure” in parallel with high-stakes decision-making and manual skills without risking patient harm is medical simulation. This approach tries to replicate real-world clinical scenarios using standardized patients (actors) or high-fidelity manikins [8]. While medical simulation has emerged as a strategy to allow for “safe failure,” its application is predominantly limited to clinical skills and diagnostic scenarios. This safety, however, is often one-sided: while the "patient" is protected, the student's psychological safety is frequently at risk [9]. Moreover, as simulation is often used as a high-stakes evaluative tool, performing under observation can, paradoxically, lead students to significant stress and a "fear of failure [9]". Finally, simulation-based training is often constrained by significant resource requirements, including specialized high-cost equipment, dedicated physical spaces, and intensive faculty supervision [8]. These limitations make simulation a distant framework for naturally learning broader socio-technical competencies around new technologies, especially in Latin America, where its use is restrained due to the economic resources of the medical colleges [3].

Productive Failure in Engineering Education

In contrast to medical education, Engineering Education has a long-standing tradition of embracing failure as a foundational epistemic resource. Productive Failure theorizes that engaging students in complex, ill-structured problems *before* receiving direct instruction leads to deeper conceptual understanding [4]. In a Productive Failure design, the initial inability to solve the problem is not a deficit, but a mechanism that activates prior knowledge and highlights the gaps that new tools must fill.

The transition from experiencing Productive Failure to achieving conceptual mastery is mediated by the formal structures of Design Theory and Iteration Theory. Design Theory provides an overarching epistemological framework for solving ill-defined problems, focusing on the synthesis of technical constraints and creative exploration to meet specific needs [11]. Within this framework, Iteration Theory defines the recursive process where each cycle of trial, feedback, and modification informs the next version of a solution [12]. Together, these theories provide how engineering students manage the inherent frustration of failure by transforming their initial inability to solve a problem into a systematic inquiry [5].

Both Iteration Theory and Design Theory explain that a system seldom works flawlessly on its first try. And for that reason, the designer prototypes, tests, detects flaws, and improves the model [5], [11]. These engineering frameworks provide a structured methodology for navigating ambiguity and complexity, transforming 'failure' from a negative outcome into a critical data point for improvement [13], [14]. This shift in perspective is instrumental in motivating the embracement of new technologies and fostering autonomous experimentation.

By framing technological engagement as an iterative design process rather than a linear search for information, students are granted psychological and procedural permission to explore technological tools without the fear of immediate 'incorrectness.' This engineering-based mindset contrasts medical mindset, and encourages learners to move beyond passive consumption, empowering them to take technical agency over the tool [3]. Consequently, the adoption of new technologies becomes less about mastering a static set of rules and more about developing a resilient, self-directed habit of trial, verification, and refinement—skills that are essential in the rapidly evolving landscape of modern professional practice [14] and potentially valuable for medical education training.

Methods

Research Design

This study employs an exploratory qualitative research design to identify emerging patterns of technical behavior among non-engineering medical student cohorts. Given the emerging nature of "prompt engineering" as a cross-disciplinary skill, an exploratory approach allows for the generation of hypotheses regarding how students transfer domain knowledge into technical constraints.

We employed qualitative thematic analysis to analyze our data to identify, analyze, and categorize patterns, synthesizing those into themes [15]. This method was selected to systematically categorize students' cognitive strategies and map them against the phases of failure and recovery, as well as the posterior reflection on the activity.

Participants and Context

The study was conducted within a Biological Evolution course at the College of Medicine of the Antonio Nariño University in Colombia, South America. The participants (N=5) were medical students spanning different levels of academic maturity as shown in Table 1. All participants reported previous informal use of AI tools but had no formal technical training in prompt engineering or AI architecture.

Table 1. Participant Demographics and Experience Profile

Participant ID	Gender	Age	Semester	Academic Stage	Previous AI Use
Student 1	Male	18	II	Novice	Yes (Informal)
Student 2	Male	18	II	Novice	Yes (Informal)
Student 3	Female	18	I	Novice	Yes (Informal)
Student 4	Female	20	VI	Intermediate	Yes (Informal)
Student 5	Female	22	VI	Intermediate	Yes (Informal)

Rationale for Sample Size

This study adopts a methodological focus on thick description [16]. A purposive sample of 5 participants was used to explore iterative behaviors and reflections of medical students after the purpose activity. By centering the analysis on information power, the research prioritizes the richness and context of the narrative, which is essential for identifying the underlying mechanisms of student engagement [15].

We engaged with the following rationale for this sample size: the inclusion of both early-stage (novice) and mid-stage (intermediate) students allowed variations in reflections across different stages of academic formation. Given that the "unguided" nature of the task triggered personal emotions and unique "Productive Failure" cycles, a smaller sample allowed for a granular analysis of individual cognitive shifts [5].

The Unguided Task

Students were tasked with using their preferred GAI tools (such as ChatGPT or DALL-E) to create visual conceptual representations of various Creation Myths, which served as the sole visual aid for a class exposition. The pedagogical core of the task required these images to be precise enough to explain the myths to their peers; therefore, students had to ensure that the final outputs accurately reflected specific cultural and narrative details, such as geography, vestiary, ethnicity, and distinct mythological characteristics. Students were expected to achieve this level of detail without any prior instruction on prompt engineering, parameter tuning, or algorithmic limitations.

Data Collection, Analysis, and Trustworthiness

We collected data via semi-structured interviews lasting 60 minutes, conducted following the activity. The interview protocol focused on three core dimensions of the student experience. First, the procedural aspect of the interaction, in which we asked participants to describe how they formulated, tested, and refined their prompts in response to the system's output. Second, the epistemic audit in which we asked how they developed verification strategies and how encountering algorithmic bias influenced their reflection. Finally, we asked about ethical use and their reflexivity regarding AI in medical training. All interviews were transcribed verbatim by a research assistant. Transcripts were subsequently translated into English.

We analyzed the transcripts using a hybrid inductive-deductive approach [15]. The author who conducted the interviews reviewed the audio recordings alongside the transcripts to generate an initial set of broad, descriptive codes. This preliminary coding phase was inductively informed by familiarity with participants' narratives, identifying raw actions such as "trying again," "checking articles," or expressions of "annoyance."

Independently, the second author conducted a manual, line-by-line coding process in Excel, relying solely on the transcriptions, and generated a separate set of emergent codes grounded directly in participants' language. Following the independent coding, we compared their coding structures, identified areas of convergence and divergence, and refined codes through iterative consensus discussions.

In a second analytic phase, inductively derived codes were mapped deductively onto theoretical constructs aligned with Productive Failure and Iterative Theory (iteration and validation). This allowed us to interpret procedural student behaviors using established engineering education frameworks and connect them to the context of medical education.

To ensure the trustworthiness and rigor of the findings, several strategies were employed in line with Lincoln and Guba's criteria [17]. Credibility was strengthened through independent dual coding and consensus between authors with varying degrees of familiarity with the data. Dependability was supported by maintaining an audit trail, including versions of the codebook and a thematic matrix with operational definitions and illustrative excerpts. Confirmability was addressed through reflexive dialogue between authors from engineering education and medical education backgrounds. No member checking was conducted; therefore, trustworthiness relied primarily on independent coding, consensus procedures, and systematic documentation of analytic decisions.

Results

Our thematic analysis revealed five major categories presented in Table 2. Despite the lack of formal training, participants spontaneously developed heuristics for managing system ambiguity, which we categorize into five levels of analysis.

Table 2. Thematic Analysis Categories for Students' Exploration

Analytical Category	Thematic Code	Operational Definition	Key Evidence (Translated Excerpts)
I. Technological Agency	*Iterative Prompt Refinement*	Cyclic process of adjusting instructions based on visual feedback to meet a specific design goal.	"It was a process of iterating trial and error... I gave more details each time to adjust the image." (S1)
	Co-creational Logic	Perception of AI as a "trainable" entity that requires expert human direction and technical logic.	"It's like educating it... if you speak to it in its own way [algorithmic logic], it will understand better." (S2)

	Peer Literacy Variance	Identification of unequal proficiency and diverse usage patterns among classmates, highlighting the lack of a common technical baseline.	“It’s worrying; only 30% know how to use AI correctly... others barely know what a prompt is.” (S1)
II. Ethical introspection	*Algorithmic Bias Identification*	Critical detection of racial, cultural, or technical biases and stereotypes in AI-generated outputs.	“... I asked for Colombian indigenous people, and it gave me people from India with Hindu clothing.” (S2)
	Cognitive Erosion Anxiety	Introspective fear of “mental laziness” and the loss of clinical/critical skills due to technology dependency.	“The brain becomes lazy... if one graduates based solely on AI, they would be a mediocre professional.” (S4)
	Epistemic Responsibility	Ethical recognition that human validation and prior knowledge are mandatory for knowledge safety.	“If I had told the application to generate it without prior knowledge, I wouldn’t have known what was correct.” (S5)
III. Epistemic Validation	*Student Auditor*	Active verification of AI-generated content against academic literature to identify hallucinations or errors.	“Sometimes you ask [the AI] and it answers, so I say, 'well, I’m going to go look somewhere else... You start looking for pages about the topic, and then you start comparing information. And on many occasions, it does happen that suddenly the information in artificial intelligence is... a bit more limited.” (S3)
	Epistemic safeguarding	Use of prior domain knowledge as a protective filter to prevent the internalization of incorrect AI information.	“Directly, if I had told the application to generate it without prior knowledge, I wouldn’t have known what was correct and what was not.” (S5)
IV. Socio-Affective Dynamics	*Technological Wonder*	Positive affective states that drive the initial engagement and sustain interest during unguided exploration.	“To be honest, I had never used [AI] to create images before; so, I was truly impressed by how the AI could create anything one could imagine.” (S2)
	Functional Annoyance	A low-stakes negative emotion arising from system errors that,	“At first, it was a bit annoying, but more than annoying, it was actually

		instead of causing paralysis, triggers further iterative attempts.	amusing, so to speak, because as I said, the [AI] is very stereotypical. ...But as I began highlighting specific details here, it started adding more detail little by little each time.” (S1)
V. Institutional Friction	*Institutional Resistance*	Conflict between the students' use of emerging technology and perceived academic “traditionalism.”	“Medicine remains very old-fashioned... professors learned that way, so they see it as the right way.” (S5)
	De-stigmatization of AI	Transition from the hidden or “taboo” to the recognition of AI as a legitimate professional tool.	“It [GAI] went from being something very hidden, to something a bit more visible... to stop doing something in secret.” (S1)

1. *Technological Agency*

Students described their interaction with GAI not as a simple query-response mechanism, but as a recursive design process triggered by initial failure. This aligns with Productive Failure, where the initial struggle serves as a generative mechanism for learning. Rather than passively accepting the first output, students engaged in a cycle of testing and refinement. It also aligns with Design Theory, where the initial “prompt” serves as a prototype that must be refined through successive iterations.

- S1*: “It was a process of iterating trial and error... I gave more details each time to adjust the image.”*
- S2: *“It’s like educating it... If you speak to it in its own way [algorithmic logic], it will understand better.”*

This “education” of the model suggests that students recognized the necessity of understanding the system’s underlying logic to achieve a desired output, a core competency in engineering literacy.

Simultaneously, students recognized a digital divide among their own cohort. Student 1 (Semester II) estimated a low level of true AI literacy among peers:

- S1: *“It’s worrying; only 30% know how to use AI correctly... others barely know what a prompt is.”*

2. *Ethical Introspection*

Students engaged in spontaneous ethical auditing when faced with the tool’s limitations.

- S2 encountered a direct instance of dataset bias: *“... I asked for Colombian indigenous people, and it gave me people from India with Hindu clothing.”*

This failure forced the student to confront the socio-technical reality that algorithms are not neutral. Furthermore, this technical skepticism extended to reflections about professional practice. Student 4 reflected on the long-term impact of reliance on automation:

- S4: *“The brain becomes lazy... if one graduates based solely on AI, they would be a mediocre professional.”*

These reflections indicate that unguided interaction forces students to confront the biases of AI, fostering an early sense of professional responsibility.

3. *Epistemic Validation*

A critical finding was the students' insistence on domain-specific knowledge as a prerequisite for using AI. They viewed themselves as "auditors" of the tool, demonstrating that unguided use triggered a validation loop.

- S5 highlighted the danger of using the tool without domain expertise*:* "*Directly, if I had told the application to generate it without prior knowledge, I wouldn't have known what was correct and what was not.*"

The student recognized that human auditory is essential for quality control. The AI generates raw material, but the student must possess the "ground truth" to validate it.

4. *Socio-Affective Dynamics*

Another significant finding was the emergence of a resilient affective state characterized by "functional annoyance" and humor, rather than cognitive paralysis. These events happened when students were confronted with AI platforms' delays or failures in addressing their requests. In those cases, students viewed system hallucinations as technical challenges to be overcome through iteration.

- S1 described this shift from frustration to iterative action: "*At first, it was a bit annoying, but more than annoying, it was actually amusing [gracioso]... because the AI is very stereotypical… But as I began highlighting specific details... it started adding more detail little by little each time.*"

5. *Institutional Friction*

Finally, the analysis identified a systemic gap between students' innovation and desire to engage in new technologies and institutional traditionalism.

- S1: *"It [use of GAI] went from being something very hidden, to something a bit more visible... to stop doing something in secret."*
- S5: *"Medicine remains very old-fashioned... professors learned that way, so they see it as the right way."*

Discussion

The results of this study suggest that the practice of "unguided prompting" among medical students could be reframed within the Productive Failure framework and Iteration Theory. Rather than a risk to academic integrity, the GAI use, nowadays a habitual process among students, represents an unintentional but valuable design process in which, by removing direct instruction, they are forced to develop their own interaction protocols and effective technical competence that emerges after multiple iterative cycles.

These finding counters the narrative that GAI promotes student passivity [2]. On the contrary, the hidden nature of GAI elevated students' roles to those of auditors. A surprising finding was the affective climate that arose during the activity, which contrasts with the traditional medical education approach. In high-stakes medical simulation environments, while "failure" is allowed, it is often accompanied by high levels of stress and a punitive psychological weight due to the potential risk to patient safety. In our study, however, legitimizing GAI within a classroom exercise during the first years of the program replaced the fear of error with wonder and curiosity.

This shift created a safe “low stakes” environment where students felt no need to hide their use of the tool or their errors during the exercise, and instead of paralyzing frustration, system errors were perceived merely as a functional “annoyance.” This specific affective state acted as a technical engine; the students’ annoyance with unsatisfactory results did not lead to abandonment or anxiety but rather triggered a recursive process of inquiry and refinement. This dynamic shows a transition from passive consumers of technology to critical “designers” of AI-mediated knowledge that attempt to align their natural language with the system’s computational logic. This interaction demonstrates a nascent form of engineering literacy, a mindset typically absent in the “no-error” paradigm of traditional medical curricula.

Furthermore, this engagement fosters an early sense of professional responsibility, a core component of any profession’s ethics. By spontaneously identifying algorithmic biases and expressing a fear of “professional mediocrity,” students demonstrated ethical introspection. They recognized that technical tools could produce flawed results if not managed with personal and professional integrity. And by acting as “auditors,” they embraced a human responsibility in the process of creation, recognizing that system outputs must be rigorously approved or rejected on disciplinary grounds.

Despite this student-led AI exploration, a critical tension remains with institutional medical “traditionalism.” While students navigate the complexity of GAI through engineering-like iteration, the broader socio-technical system often creates friction through rigid pedagogical standards. This highlights the importance of interdisciplinary studies, such as ours, to overcome these barriers.

Implications for Engineering Education

Studying the technological trajectories of medical students provides engineering educators with a “disciplinary mirror” that validates the transversality of engineering education approaches. Our research shows how frameworks like Productive Failure in the Iterative Theory context are fundamental cognitive strategies applicable to areas such as medical education. Productive failure can position students in a resilient affective state driven by “functional annoyance” and wonder. This suggests that technical persistence is fueled by an emotional engagement with the tool, allowing engineering educators to better understand how the design of low-stakes environments can catalyze spontaneous iteration without the need for constant, direct instruction. Furthermore, this research provides an alternative perspective to the core assumption that technical agency in AI necessarily requires mastery of formal syntax [1], indicating that design thinking can emerge directly from domain expertise and natural language interaction.

The most profound implication of this work lies in the untapped potential of interdisciplinary collaboration between medical education and engineering education. Frameworks considered foundational in engineering, such as Iteration and Design Theory, become revolutionary tools when applied to medical training. This study demonstrates that what engineering education takes for granted can serve as a transformative resource for the human-centered sciences, offering a new scope for developing technical resilience in future physicians.

Limitations of the study

A primary limitation of this study is the sample size; however, this was a deliberate choice to prioritize “information power” and “thick description” [16] of the participants’ iterative processes. Additionally, digital artifacts such as prompt logs or final images were not collected, as the research highlights the students’ reflections on technical fidelity. Future studies could expand upon these findings by incorporating larger cohorts and analyzing digital artifacts to provide a more comprehensive validation of emergent learning trajectories in non-engineering students compared with those in engineering, in the United States and countries of the Global South.

Conclusions

This study demonstrates that the unguided use of Generative AI triggers complex learning processes in non-engineering students that align closely with established engineering education frameworks. Through the lens of Productive Failure and Iteration Theory, we observed that students did not merely consume information: they developed validation strategies, ethical critical thinking, and technical agency.

We conclude that the "failure" to provide initial guidance was a catalyst for learning, as it forced students to bridge the gap between their domain knowledge and the tool's algorithmic logic. By embracing the iterative, error-tolerant mindset of engineering, rather than the "no-error" paradigm of traditional medicine, professional education can better prepare future practitioners to navigate the complexities of an AI-driven landscape without compromising their clinical rigor or ethical responsibility.

On the other hand, by observing that iteration flourishes authentically in unguided spaces, this work gives a different approach to the over-structuring of prompt design in traditional curricula that is being translated to STEM education. Ultimately, this interdisciplinary perspective proposes a shift where technological autonomy fosters genuine technical resilience. Finally, the intersection of medical and engineering education can inform and transform both disciplines. We suggest that future pedagogical efforts continue to explore this interdisciplinary dialogue, building a shared educational landscape.

References

[1] C. Tran, B. N. Hryciw, S. W. Moore, A. Chaput, and A. J. E. Seely, "Perceptions and Use of Generative Artificial Intelligence in Medical Students: A Multicenter Survey," *J. Med. Educ. Curric. Dev.*, vol. 12, Oct. 2025, doi: 10.1177/23821205251391969.

[2] L. McCoy, N. Ganesan, V. Rajagopalan, D. McKell, D. F. Niño, and M. C. Swaim, "A Training Needs Analysis for AI and Generative AI in Medical Education: Perspectives of Faculty and Students," *J. Med. Educ. Curric. Dev.*, vol. 12, Dec. 2025, doi: 10.1177/23821205251339226.

[3] N. Komasawa and M. Yokohira, "Generative Artificial Intelligence (AI) in Medical Education: A Narrative Review of the Challenges and Possibilities for Future Professionalism," *Cureus*, Jun. 2025, doi: 10.7759/cureus.86316.

[4] M. Kapur, "Productive failure," *Cogn. Instr.*, vol. 26, no. 3, pp. 379–424, Jul. 2008, doi: 10.1080/07370000802212669.

[5] S. Purzer, R. Jiang, I. Lyss-Loren, F. Demirci, and J. Quintana-Cifuentes, "Reasoning through iteration: How do engineering design projects promote student learning and self-efficacy?"

[6] J. Sandars and P.-S. Goh, "Design Thinking in Medical Education: The Key Features and Practical Application," *J. Med. Educ. Curric. Dev.*, vol. 7, Jan. 2020, doi: 10.1177/2382120520926518.

[7] J. Cox, S. Grover, S. Kharileh, K. Haber, N. Savage, and J. Fuller, "Consequences of Academic Failure in Medical School: The Student's Perspective," *Cureus*, Jul. 2025, doi: 10.7759/cureus.87996.

[8] C. Elendu *et al.*, “The impact of simulation-based training in medical education: A review,” Jul. 05, 2024, *Lippincott Williams and Wilkins*. doi: 10.1097/MD.0000000000038813.

[9] G. M. Bump and F. P. Cladis, “Psychological Safety in Medical Education, Another Challenge to Tackle?,” *J. Gen. Intern. Med.*, vol. 40, no. 1, pp. 41–45, Jan. 2025, doi: 10.1007/s11606-024-09166-y.

[10] A. V Potharazu *et al.*, “Design thinking in North American undergraduate medical education,” *The Journal of Health Design*, vol. 9, no. 2, pp. 663–674, Dec. 2024, doi: 10.21853/jhd.2024.233.

[11] M. Gottlieb, E. Wagner, A. Wagner, and T. Chan, “Applying Design Thinking Principles to Curricular Development in Medical Education,” Jan. 01, 2017, *John Wiley and Sons Inc*. doi: 10.1002/aet2.10003.

[12] W. D. Moscoso-Barrera, “An Iterative Design Approach in Biomedical Engineering Student Group Projects.”

[13] R. S. Adams, “Understanding design iteration: representations from an empirical study,” in *Common Ground - DRS International Conference 2002*, Design Research Society, Sep. 2002. doi: 10.21606/drs.2002.12.

[14] S. Quon, S. Thompson, D. Nguyen, E. Carter, and M. Lee, “Integrating engineering into medical education: A scoping review,” *Curr. Med. Res. Pract.*, vol. 15, no. 3, pp. 100–104, May 2025, doi: 10.4103/cmrp.cmrp_60_25.

[15] S. K. Ahmed *et al.*, “Using thematic analysis in qualitative research,” *Journal of Medicine, Surgery, and Public Health*, vol. 6, p. 100198, Aug. 2025, doi: 10.1016/j.glmedi.2025.100198.

[16] C. Geertz, *The Interpretation of Cultures Selected Essays*. 1973.

[17] Y. S. Lincoln and E. G. Guba, *Naturalistic Inquiry*. Sage Publications, 1985.