Publications of the Astronomical Society of Australia (2020), 1–21
doi:

RESEARCH PAPER

Revolutionizing Intelligent Tutoring Systems: A Student-Centric Approach to Personalized Learning Using Bloom Taxonomy

Ayoub Sbai[1] and Wadie El Amrani[2]

[1]Systems Engineer
[2]AI Researcher

(Received ; revised ; accepted ; first published online)

Abstract

The lack of personalization in education hampers struggling students' ability to address learning gaps, contributing to burnout, high dropout rates, and absenteeism. This issue arises from misguided approaches to personalization that focus on poor learning proxies, misdiagnosis of causes-and-effects and an unhealthy emphasis on multi-modal course content, as current approaches to Intelligent Tutoring Systems (ITS) often emphasize domain-specific expertise (e.g., math or coding), which neglects the core bottleneck in education: 1- lack of a good student model 2- lack of a learning framework to utilize the student model.

Our solution addresses this by combining both elements: democratizing access to an accurate student model and applying Bloom Taxonomy as a learning framework to utilize it. This contrarian approach personalizes the student profile rather than the domain, leveraging Bloom Taxonomy to build student-centric ITS. Key to our innovation is delegating student-model expertise to large language models (LLMs) while leaving domain expertise to students and teachers, thus ensuring both personalization and autonomy. This avoids issues like LLM hallucinations in domain-specific contexts which can worsen student performance and enhances accuracy in student modeling. Unlike BloomBERT, Khanmigo, and TahseenAI, which face limitations from narrow focuses or flawed assumptions about learning, our solution utilizes learning's first principles to efficiently exploit student models. By focusing on the foundational student model, we aim to address global educational challenges, including those highlighted by Morocco's low PISA ranking, and to extend EdTech beyond K-12 and language learning markets.

Keywords: Bloom Taxonomy, Bloom's 2 sigma problem, Mastery Learning, VARK model, Hypercorrection Effect, Cognitive Tutors, Intelligent Tutoring System, Cognitivism, Connectivism, Large Language Models, Artificial Intelligence

1. INTRODUCTION

Education is the process of acquiring knowledge, skills, values, and attitudes through various formal and informal methods. It is a lifelong endeavor that begins from an early age and continues throughout one's life. Education encompasses a wide range of activities, including teaching, learning, training, and research. Education holds immense value in our lives, much more than usually believed. It empowers individuals by providing knowledge, skills, and critical thinking abilities (cognitive domain) as well as civic engagement and personal development (affective domain). Education opens doors to economic opportunities, reducing poverty and promoting social mobility. It fosters personal growth, equips individuals with problem-solving skills, and encourages responsible citizenship. Education contributes to improved health outcomes, cultural preservation, and appreciation. But most importantly, It also drives innovation and technological advancement that is responsible for bootstrapping the state of the world to the next level in an almost virtually-free civilizational growth. Ultimately, education has the power to transform lives, strengthen communities, and create a better future for us all. It's like a sleight of hand magic trick that the human species are evolutionarily lucky to possess, right? So why fumble on it? Axiologically speaking, education has an intrinsic value to the world, something that is and should be valued for what it inherently is. We can't afford the cultural, social and civilizational costs incurred due to the value misalignment that result from viewing education as instrumental. It never is and it should never be. Solving educational problems must be the core focus of all problems. That's due to Thiel's law, we must deal with the foundation of all problems before going on and solving those subproblems (that are usually imbued with vested interest). I hypothesize that most, if not all, of problems boil down to educational problems from the core.

The history of education is a fascinating journey that has witnessed significant milestones and transformations. It all began with the invention of writing, which revolutionized the way knowledge was recorded and transmitted. With writing, information could be accumulated and preserved across generations, paving the way for the emergence of formal education systems. In ancient times, education was often an exclusive privilege of the elite. The ancient Greek elites, for instance, had access to renowned institutions like Plato's Academy, where students received instruction in a range of subjects. The Great Library of Alexandria also played a pivotal role as a center of knowledge and learning. The invention of the printing press in the 15th century brought about a revolutionary change. Johannes Gutenberg's creation made books more accessible and affordable, democratizing education. This breakthrough propelled the spread of literacy and opened new doors for the dissemination of knowledge. During the Middle Ages, education was predominantly limited to private institutions, religious establishments, and individual tutors. Religious studies and Latin instruction were central to the curriculum. However, this era laid the groundwork for educational developments that would follow. The 18th and 19th centuries marked the introduction of public education systems. These systems sought to provide education to the general population and promoted concepts like compulsory education and free public schools. This shift aimed to ensure that education was accessible to all members of society, regardless of social status or wealth. As the 20th century progressed, a trend towards standardization emerged. Curricula were standardized, and efforts were made to establish uniform methods of assessing student performance through standardized tests and evaluation that would latter have their own kind of perverse incentives. Teaching certification standards were also introduced to ensure quality education delivery. The latter part of the 20th century and the 21st century witnessed the rise of education technologies and online education. Cognitive tutors, Learning Management Systems (LMS), and Human-Computer Interaction (HCI) technologies played crucial roles in enhancing educational experiences. The advent of Massive Open Online Courses (MOOCs) enabled widespread access to high-quality education globally. Moreover, the emergence of e-learning as well as the availability of resources on the internet has revived the spirit of autodidactism. In this era of advanced technology, language models and chatbots have emerged as tools to enhance education further in the most unprecedented of manners. They offer personalized learning experiences, answer students' questions, provide explanations, and support learners in various subjects. Language models can assist with language learning and offer writing assistance, while chatbots can deliver educational resources tailored to individual needs, something that this paper tries to take advantage of in the form of a proposed solution to the research problem. The history of education showcases an ongoing evolution, driven by societal changes, technological advancements, and a growing understanding of effective teaching and learning methods. Each milestone has shaped the education landscape, aiming to improve access, quality, and outcomes for learners worldwide.

As previously mentioned, education is a crucial aspect of society, shaping the future of individuals and communities. However, several significant challenges persist within the educational landscape. One of the pressing issues is the achievement gap, which refers to the disparities in academic performance and educational outcomes among different groups of students, often linked to socioeconomic factors and unequal access to resources. The dropout rate is another concern, reflecting students who prematurely leave school without completing their education, which can hinder their future prospects. Additionally, the integration and access to technology in education is crucial, as it can enhance learning opportunities (which is the summum bonum of this project), but disparities in access can perpetuate inequalities. The relevance of the curriculum is vital for engaging students and preparing them for the real world. Ensuring educational access and equity in education is essential, as marginalized and underserved populations often face barriers in receiving quality education. The overall quality of education is another area of focus, encompassing effective teaching practices, supportive learning environments, and up-to-date resources. Adequate funding and resource allocation are crucial to provide equitable educational opportunities for all students. Proper teacher recruitment and evaluation strategies are necessary to ensure a qualified and diverse teaching workforce that can meet students' needs. Student engagement is essential for fostering a love of learning and maximizing educational outcomes. Maintaining accountability and assessment measures helps monitor progress, ensure educational standards, and identify areas for improvement. The provision of special education services for students with disabilities is vital to support their unique needs and promote inclusivity. Preparing students for college and career readiness is a critical goal, equipping them with the skills and knowledge necessary for success beyond their educational journey. Lastly, aligning education with the needs of the workforce is crucial to bridge the gap between education and employment opportunities. Addressing these multifaceted challenges requires collaboration, innovation, and a commitment to providing every student with an equitable and high-quality education. This paper is more oriented towards technology integration and access as well as educational equity and quality maintenance manifested in closing the unnecessary achievement gap. Funding is not enough. It's primarily redistributive and only secondarily productive. Funding can only do so much up to a value bound, its linearly scalable (the bigger the problem → the bigger the funding needed → the bigger the cost) and thus intrinsically costful. Moreover, funding doesn't guarantee curing the ''metalearning diseases" and bad habits swirling around today's classrooms akin to a sort of cognitive epidemic where these helpless and often harmful ''metalearning genes" hosts on the students' minds in the form of parasites. As a reminder, the study unit of this project is the ''student". Funding does little to nothing to enhance their cognitive faculty, at least not directly (cognitive problems require cognitive solutions that directly target each and every student on an individual level) and therefore it's not a ''cognitive vaccine". This project tries to ''do more with less" by affording the scalability marginal cost (which is zero in this case). I hypothesize that the solution to educational problems are: metalearning techniques proper acquisition, passion enculturation and pedagogy quality. Now, due to funding's intrinsic non-zero marginal cost, no matter how good pedagogy quality is, it won't scale cost-free to all students. The problem has to be solved organically and that's what this paper is for; automating/personalizing metalearning techniques both in encoding phase and retreival phase with network effect-driven active collaboration (could be used in blended learning during the lecture; something that GPTs on their own can't afford nor are fine-tuned to)

The purpose of this paper is to examine the issues surrounding the learning experience in engineering schools in various situations restricted to three specific areas of concern namely; lecture quality, TD session and test assessment. The focus on these measures is justified by previous literature on the matter. It is practically worthy to frame students' learning experiences as a twofold phenomenon with two phases; the encoding phase and the retrieval phase. In the case of lecture quality, it is conspicuously playing a large part of the student information processing, actively or otherwise, lecture environment has been proven essential to educational performance, be it success or failure (Smith, 1979; Lang, 1996)[1]. On another aisle, school inputs; examples like lecture hall design, lecture/students ratio and library experience have been proven to contribute to student's academic success. As matter of empirical fact, several components of lecture hall are chiefly affecting the learning experience (Fleming, 1999)[2], the paragon of these characteristics is the lecture size, shape, acoustic quality/noise control (Lang, 1996; Bligh, 1972)[3]. Adding to that, students/teachers ratio correlates significantly negatively with students achievement (at -.561) (Koc, 2015; Ajani and Akinyele, 2014)[4]. This by and large is explained by the trade-off between the amount of instructions demanded during class and individual attentions from teachers (Bayo, 2005)[5] where small classes, in contrast to large lecture halls, have been proven effective - proportionally with the longevity- in enhancing academic achievements. Similar results have been documented in the famous Bloom's 2 sigma problem (Bloom, 1982)[6] suggesting that the average student tutored one-to-one using mastery learning techniques (Bloom, 1968)[7] performed two standard deviations better than students educated in a classroom environment with one teacher to 30 students, with or without mastery learning (more precisely,the average tutored student was above 98 percent of the students in the control class. Moreover, about 90 percent of the tutored students attained the level of summative achievement reached by only the highest 20 percent" of the control class). Briefly, Bloom's 2 sigma problem focus on two fundamental parameters or variables: student size (thus teacher/individual student attention ratio via comparing individual attention to students versus group performance assessment, also non-individualized learning pace for each student as well as averaging out different learning rates by methylating unsystematic factors resulting from non-shared environmental factors (Plomin and Daniels ''Gloomy Prospect", 1987)[8] –like interest, background knowledge, prerequisite level, ability, which goes against the principles of Mastery Learning (Bloom, 1968) that sets time required for student's learning curve of achieving the same level of mastery as its most important parameter) and instructional method. However, due to economic pressure epitomized in the trade-off between the school funding and its need to graduate more students, lecture halls are not going anywhere any soon. On a par with the previous, library usage is no exception. In fact, questioning the role of libraries in enhancing the accumulative process of information gathering has been described as "heretical" by George Kuh and Robert Gonyea. It is an empirical fact that library usage, in whatever form, has a plenary bearing on student retention and academic success where it's been reported that students who use library scored higher GPA compared to the otherwise - regardless of demographic characteristics, background knowledge (Soria, 2013)[9]. In the wake of all that, it is deemed foolishly naive to turn a blind eye to the significance of school inputs in dictating the quality of academic performance. Our schools today, especially coming under the pressure of a far-reaching cognitive advancement in the students mental artillery wrought by an increasingly intellectually stimulating, cognitively demanding environment (Flynn, 1981)[10], can't afford the mismatch between supply and demand of today's pupils through committing the "Coleman Reports Fallacy". In a technologically rich environment such as ours today, our ignorance of its potentials in galvanizing our educational system is becoming a hotbed that continues to rear its ugly head by having our heads below the parapet, a shock therapy is surely needed and it can only come at the detriment of reluctance to make a change for the better, now that data-driven cognitive tutors are here thanks to the recent development in Natural Language Processing technology (from the introduction of LSTM to Transfomers). Up until now, the focus has been directed on the bottlenecks that spearhead the encoding phase. It is now time to shift the focus to the retrieval phase which is - contrary to popular believes - is as important as the former if not more (Karpick, 2011)[11]. Our retrieval journey starts with TD sessions, it has long been held that "who does the work, does the learning"-taken at empirical face value, it contains some truth. In the light of the overarching importance of active retrieval phase in cementing knowledge in the minds of students (Karpick, 2011), it is crucial that we scrutinize what happens during the TD session for two reasons. First, TD sessions are usually clustered and subsequently tend to have balanced teacher/students ratio which is favorable (Koc, 2015)[12]. Second, TD sessions are almost always undertaken in relatively small classrooms which are more engaging in terms of information processing, attention and participation of students (as measured by Engaging Learning Index) (Bolden, 2017)[13]. A possible caveat of the latter claim is that engagement is a poor proxy for learning (Coe, 2015)[14] or what Richard Mayer calls "Constructivist Teacher Fallacy". However, though engagement per se may not lead to learning, it's a requirement that the possibility of learning hinges on; specifically cognitive engagement (Pickering, 2017)[15]. Therefore, even if, ideally, we take encoding phase for granted, we can't afford having TD session retrieval phase as a fly in the ointment. Fortunately, as mentioned before, data-driven metacognitive cognitive tutors are the perfect match to help with retrieval, active recall practices and metacognitive regulation. The desperate attempt to tackle all the previous issues has as its summum bonum ensuring that students have learned something; skills acquisition, critical thinking..etc, which is usually called Information Literacy, also ensuring the effectiveness and/or the quality of teaching methods. This is plausibly possible with the help of standardized tests under just-right-conditions. But what is exactly a test and why is it useful? It all started within scientific management, or Taylorism as usually referred to, and specifically the ingenious idea of using statistical indicators to evaluate and control quality which is known as quality control in business. It was Ralph Tyler who first coined the term "educational evaluation" which served as a landmark in education sector, thus the standardized tests. That said, standardized tests are objective standardized measures for assessing the learning experience of students ie; ideally answering the question of whether students learned something or no, tests are far more importantly used as diagnosis to help identify the weakness

spots of students or teaching methods so that changes take place in an evidence-based manner. Unfortunately, this is usually not the case, in fact, and in the light of the importance of tests in affecting not only the learning experience of students but the whole pedagogical system and schools more broadly (Herman, 2005)[16], tests are easy to corrupt and game the educational system with.
Usually, standardized tests are depicted as anxiety-inducing and tend to eclipse the actual purpose of education ideally epitomized in a curriculum-driven instruction, in that telling, their inferential validity is biting the dust as the line between a skills acquisition-oriented test and test-oriented test has become blurred and called to question (Cannell, 1988; Linn, Graue,and Sanders, 1989; Shepard, 1989)[17] they nonetheless have an essential place in the educational system; . Speaking of their prediction power, standardized tests have been proven to be good indicators of success in graduate school and significant academic outcomes (Kuncel, 2007)[18]. However, and with the increasing accountability and rather erroneous urge to use standardized achievement tests to measure the school's as well as teacher's quality, the whole instructional process is tailored to achieving high scores on test but without actual mastery of knowledge or skills (Rothman, 1988; Popham, 2001)[19]. Such practice has become known to be called "Teaching to the test" which is either implemented in teaching the actual items of the test or "clone" items (Popham, 2001)[20] which is predominant in our educational system for regulatory reasons in contexts where tests are used for important decisions (Haney and Madaus, 1986)[21]. The caveat is, as Charles Goodhart puts it, "when a measure becomes a target, it ceases to be a good measure" and standardized tests are no exception. In fact, and even worse, tests are apt to corrupt the very construct of interest they purport to measure which is the mastery level of knowledge and learning when used for control purposes (Campbell, 1976)[22]. As matter of fact, test validity and inference are often subjects to controversies, for instance, tests may drop the ball on the instructional process, even distort the curriculum or at least alter it substantially when the situation demands it (Herman, 2005). If misused, tests can germinate and become an educational hotbed that stunts the very essence of human curiosity by, for example, trivializing the instructional process and wasting the valuable time for actual learning to take place (Bracey,1989; Dorr-Bremme and Herman, 1986)[23]. In short, the more you regulate, the less you educate. On the eve of running into a philosophical trouble of measures, it is a good practice to avoid tumbling off this epistemological turnip truck by recognizing the Is/Ought problem here; just because tests can be corrupted for regulatory and control purposes or what have you, it doesn't follow that that they are intrinsically invalid. In fact, tests are neither intrinsically valid or invalid. The validity of tests lies in how they are used, constructed and the nature of the content they cover; all of this is referred to as test preparation which is an often controversial subject of heated debates ranging on a spectrum from "legitimate" preparation; one that forbids the use of item-teaching which is using similar or, in worst case scenario, the exact standardized test items in order to practice and prepare for tests (Wrightstone, 1967)[24] to "illegitimate" preparation but more neutrally known as Criteria Referenced Instruction; which is the use of identical or similar items in both the instructional process and final assessment as to have a perfect match ensuring that the content is learned (Cohen, 1987)[25], this is also common with the idea of Measurement Driven-Instruction which contends that the instruction must eventually tailor to the items covered in the test content (Popham, 1985)[26] . And quite often than not, control is the confounder, lurking variable between validity of the test and its effect on learning process.
It is a given that the quality of learning is- and again, ideally-measured up by the quality of tests, tests are our compass in this decades-long labyrinth of education policy research and therefore dropping the ball on them is equivalent to sitting on the razor's edge. On a par with the two previous variables, LLM technology can help very massively with testing, evaluating and giving context-sensitive feedback based on cognitivist-constructivist models of learning (to be discussed later).

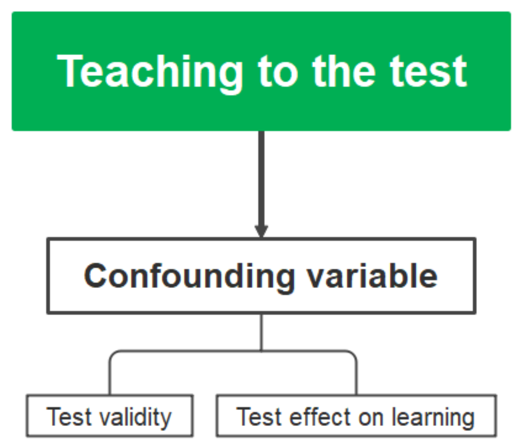

Figure 1. Illustration of the spurious relationship between test validity and its effect on learning experience

Hoping that this relatively long introduction helped clarify the reason why this paper is singling out the aforementioned criteria, it is now noteworthy to remind the reader of the outline of what comes next. This paper is structured as follow : in the first section, the light is shed on the concept of learning as well as examining various models, preferences and educational theories of learning on a par with pinpointing their relevance in the educational system. After that, section two is concerned with the methodology used in how and why the research design was chosen as such, how data was collected and it was analyzed and subsequently used. Third section is for discussing the results found whilst directing the focus on their implications and the contribution of the paper as a whole. Fourth section deals with the proposed solution to the yet-to-proven learning difficulty that students usually undergo.

2. Learning : Definition, Transfer, Styles, Proxy variables, Theories

With the bifurcation of the 20th century academic research, it is no short of a miracle that learning - as essential as it started to be depicted as and especially in an progressively intellectually-rich ethos that fostered meta-cognition (Flavell, 1979)[27] as an equally important modus operandi of exploring the world- has been the subject of a decades-long spurred debate, usually pertaining to the definition of learning, types of learning, how to transfer it, the factors involved in it, its proxy variables and so on.

2.1 *What is learning?*

Without muddying the waters while having an egg on our face out of hocus-pocus definitionalism, a simple, and I think a convincing, definition of learning would amount to the process of acquiring data from the world (values, attitudes, behaviors, ideas..etc), applying it in various situations when needed. A visually appealing explanation of learning hinges on splitting the mind in two parts; System 1 and System 2. System 1 is fast, unconscious, automatic and capacity-abundant. System 2 is the exact opposite; slow, conscious, thoughtful, analytical and capacity-constrained(7 +/- 2 items as a limit; Miller's number/law). System 1 resides in the hippocampus (Long-Term memory) while System 2 resides in the prefrontal cortex (Working/Short-Term memory) (Kahneman, 2011)[28]. Learning occurs when information is initially passed to System 2, witfully and deeply analyzed concomitant with focused attention and spaced repetition which subsequently form "chunks" ie; clustered bits of information grouped together based on how much similar they are, semantically or otherwise (G. Miller, 1959; H.Simon and W.Chase, 1973)[29]. These chunks form a bottom-up approach to the big picture (top-down) where recalling one bit of previously stored information triggers a whole series of "related" neurons to fire as to give a holistic meaning to that one bit as a part of a larger whole (neural pathways and neuroplasticity). Chunking is not a bug that averages out or abstracts "irrelevant" information, it is a computation-efficient evolutionary heuristic. These chunks are subsequently passed to system 1 as now ready at hand to be used automatically, rapidly and subconsciously when encountering similar situations that prompted the slow, conscious thinking in the first place. This process goes on recursively resulting in chunks getting bigger and bigger until the whole big picture is covered– though some would argue that top-down mountain top view is theoretically impossible to reach from an organic chunk-based approach to learning due to the underlying complexity of learning, it's like climbing a steep mountain but you first have to build the ladder from scratch along the way. That is, in the cognitive sense of the word, what we call learning.

The Information Processing theory posits that humans function as information processing systems similar to computers. It identifies three main components: information stores (memory theories), cognitive processes, and executive cognition. Unlike behaviorism, which focused solely on observable

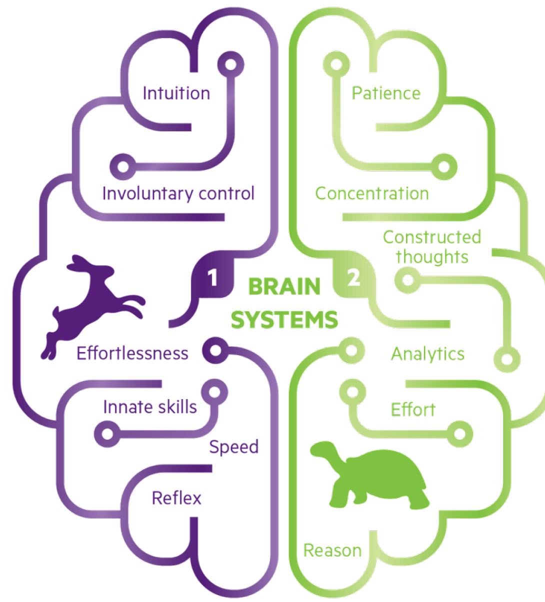

Figure 2. A visualization of System 1 and System 2

behaviors, this theory acknowledges the internal mental processes involved in learning and cognition. It also differs from Piaget's cognitive development theory, which described development as discrete stages, by suggesting a continuous development of cognitive abilities. The theory emphasizes different processing levels, ranging from shallow (structural) to deep (semantic), with memory retention being positively correlated with the depth of processing but counterintuitively independent of learning intention (T.Hyde and J. Jenkins, 1969)[30]. To achieve deep processing, principles such as elaboration, distinctiveness, personal relevance, and retrieval/application relevance are crucial. Automaticity refers to processes that occur effortlessly due to extensive practice, while overlearning involves studying beyond mere familiarity to enhance quick and easy recall.

Memory theory is a fundamental aspect of cognitive psychology, exploring how information is stored and retrieved in the human mind. Memory can be broadly categorized into several types. Sensory memory is the initial stage, encompassing iconic memory for visual stimuli and echoic memory for auditory stimuli. Short-term memory, also known as working memory, involves the temporary retention and manipulation of information. It consists of components like the phonological loop, responsible for processing verbal information, and the visuospatial sketchpad, handling visual and spatial information. Long-term memory is divided into explicit (declarative) and implicit (non-declarative) memories. Explicit memory includes semantic memory, which stores general knowledge and facts, and episodic memory, which retains personal experiences. Implicit memory encompasses priming, which influences subsequent perceptions and behavior, and procedural memory, which involves the recall of skills and habits.

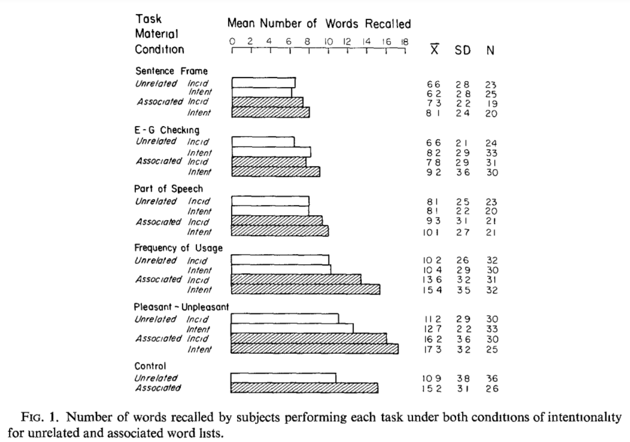

Figure 3. Positive Correlation between processing level depth and Memory retention

Two influential models shape our understanding of memory processes. The multiple stage model of memory, proposed by Atkinson and Shiffrin, suggests a progression from sensory memory to short-term memory and then to long-term memory. The working memory model, developed by Baddeley and Hitch, focuses on the active processing and maintenance of information within the working memory system.

Several laws are associated with memory. Miller's law, often referred to as Miller's magic number, suggests that the capacity of short-term memory is around seven plus or minus two items. The power law of memory, also known as the Pareto law, states that memory performance follows a pattern of a small number of items being well-remembered while the majority are forgotten.

Cognitive load theory, introduced by John Sweller, explores the amount of cognitive load or mental effort required for learning. It distinguishes between intrinsic, extraneous, and germane cognitive load. Intrinsic load relates to the inherent complexity of the task, extraneous load refers to unnecessary cognitive demands, and germane load involves the processing that contributes to meaningful learning. Cognitive load theory has implications for instructional design, including considerations of modality effect, split-attention effect, worked-examples effect, and expertise reversal effect, which impact learning outcomes.

The spacing effect, discovered by Hermann Ebbinghaus, demonstrates that spaced intervals between learning sessions result in better retention compared to massed practice. This phenomenon is often depicted by the forgetting curve, illustrating the rate at which information is forgotten over time. To optimize memory retention, active recall and spaced repetition techniques are employed, promoting the retrieval of information at intervals to enhance long-term retention.

Memory encoding format refers to how information is organized and represented in memory. Linear encoding involves sequential storage, while non-linear encoding methods, such as mind mapping, use visual and associative networks to connect concepts and aid memory retrieval.

Metamemory techniques are strategies individuals employ to enhance memory performance. Mnemonics, which involve using vivid imagery or associations, enumeration, which involves breaking information into smaller chunks or groups, and chunking, which involves grouping information into meaningful units, are effective methods for improving memory encoding and retrieval.

When talking about learning, it is a necessity to harangue the reader on the witchcraft of memory models. The lingua franca of this field is the Multiple Store model of memory (Atkinson and Shiffrin, 1968)[31], though it has its criticism that usually amount to neurological subtleties like Multiple Trace theory which is an empirically-based refinement of the former, but the overall big picture is roughly the same which is that there are three types of stages involved in the learning process with each stage associated to a specific type of memory with a specific functionality, size capacity, attention span and recall rate. When the information input is presented, the amygdala region of the brain is activated. At this stage, it's the sensory memory that runs the show except that it's too loose to handle the complexity of information presented because it's only stimulating the senses until the subject starts paying focused attention to that, and when that happens, chunks are formed after enough repetitions thanks to working memory (prefrontal cortex) which then leads to encoding the information to long-term memory.

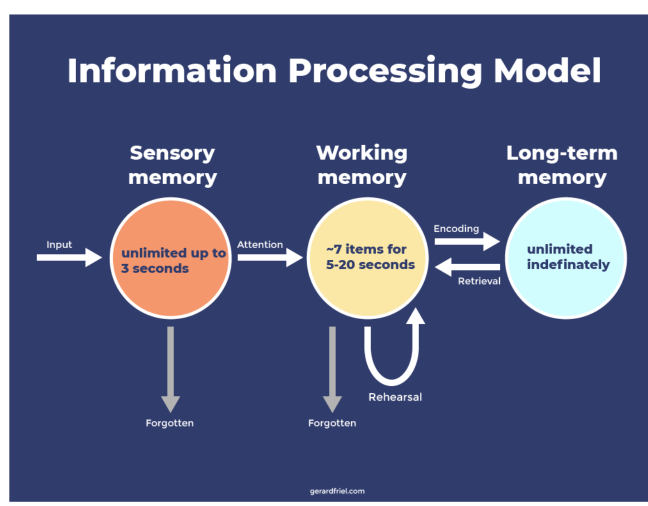

Figure 4. Multiple Store model of memory

One important phenomenon about working memory-long term memory transition is the spacing effect, as discovered by Hermann Ebbinghaus in 1885 (Ebbinghaus, 1885)[32], is a psychological phenomenon that refers to the improved retention and learning of information when it is studied and reviewed over spaced intervals of time, as opposed to studying it all at once or in a massed fashion. In simple terms, it suggests that spacing out your study or review sessions over time leads to better long-term retention compared to cramming or massed practice. The spacing effect is closely related to the concept of distributed practice, which involves spreading out learning sessions over time. Research has consistently shown that when information is studied and reviewed with intervals of time in between, the brain has a better chance to consolidate and store that information in long-term memory. This leads to improved retrieval of the information when needed in the future.

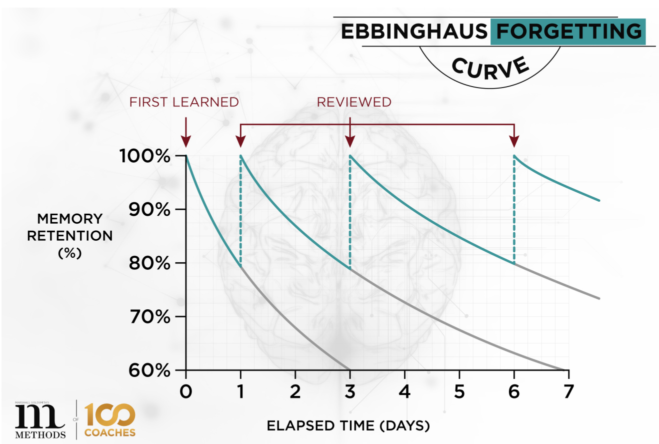

Figure 5. Forgetting curve - Spacing effect

This fundamental phenomenon is the essence of plenty of learning methods such as Active Recall and Spaced Repetition. Essential to the encoding phase is the cognitive load theory developed by John Sweller in the 1980s (John Sweller, 1980s)[33]. This would later prove to be important to the purpose of the paper as the notion of cognitive load is an essential control parameter that could, if non-redundant, enhance student's encoding process. But what is cognitive load? Cognitive load is the amount of working memory (short-term memory) resources used in the encoding phase. There are 3 types of cognitive loads that are worth mentioning; Intrinsic CT: the effort associated within a topic, Extraneous CT is more about the framing effect ie; the way information is presented to the learner, Germane CT is the work put into creating a permanent store of knowledge in long-term memory. Now, one of the tricky things about encoding or cognitive load is the redundancy problem ie; minimizing the amount of information while keeping only necessary components that are usually high-order learning. This problem has an instructional significance due to the inherent limitations of the working memory both in terms of capacity (Miller's number/law) and duration.

John Sweller suggested that the solution lies in the instructional design aspect of learning (modality effect, split-attention effect, worked-examples effect, and expertise reversal effect). Now whether this is an ideal solution or not is debatable. I would just like to mention that this problem could also be solved by abstracting it to information-theoretic terms and solving it using data compression algorithms. However, that's far from fit with the abstraction level specified for the paper. That's why, the solution is using LLM technology to do the redundancy compression for the learner so that they can engage in necessary components of learning depending on their needs (in fact, redundancy compression happens automatically once the leaner starts using the chatbot that can play the role of More Knowledgeable Other (MKO)). . In the context of this project, there are few several ways to reduce unnecessary cognitive load:

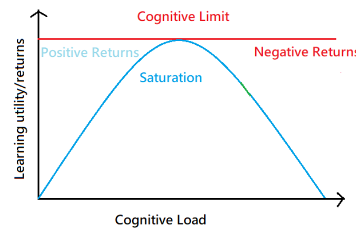

Figure 6. Cognitive Load's Laffer curve

1- Intelligent summarization: Use AI-powered language models to generate concise summaries of complex texts, reducing the cognitive load required to process large amounts of information. These summaries can highlight key concepts, main arguments, and relevant examples.

2- Adaptive learning pathways: Employ AI algorithms to personalize learning pathways based on individual learners' needs and proficiency levels. By assessing learners' prior knowledge and progress, the system can suggest appropriate learning resources and activities, reducing cognitive overload caused by irrelevant or excessively challenging material.

3- Contextualized explanations: Provide learners with contextualized explanations using AI language models. These models can generate explanations tailored to specific examples or problem-solving scenarios, helping learners grasp complex concepts and reducing the cognitive effort required to understand abstract or technical information.

4- Visualizations and interactive simulations: Utilize AI techniques to create interactive visualizations and simulations that facilitate understanding of complex concepts. These visual representations can help learners grasp abstract ideas more easily and reduce the cognitive load associated with mentally visualizing or conceptualizing complex information.

5- Concept mapping and knowledge organization: AI-powered concept mapping tools can assist learners in organizing and structuring their knowledge. These tools can help learners create hierarchical relationships between concepts, highlight connections, and identify knowledge gaps, reducing cognitive load by providing a clear mental framework for understanding and organizing information.

6- Intelligent assessment and feedback: AI can be used to provide personalized and timely feedback on learners' performance. Adaptive assessment techniques can identify areas of strengths and weaknesses, allowing learners to focus their efforts on specific areas that require improvement, reducing cognitive load by guiding learners towards targeted learning activities.

7- Natural language interaction: AI-powered chatbots or virtual assistants can engage learners in natural language conversations, answering questions, providing explanations, and offering guidance. These interactive agents can reduce cognitive load by providing on-demand support and helping learners navigate complex topics more effectively.

8- Scaffolding and guided learning: AI can provide scaffolded support to learners, breaking down complex tasks into smaller, manageable steps. This approach gradually reduces support as learners gain proficiency, reducing cognitive load by providing structured learning experiences that align with their abilities.

9- Automatic translation and language support: AI language models can assist learners in overcoming language barriers by providing automatic translation, language support, or simplifying complex texts. This reduces cognitive load for learners who may struggle with language comprehension, allowing them to focus on understanding the content rather than deciphering language nuances.

10- Adaptive time management: AI algorithms can help learners optimize their study time and schedules. By analyzing learning patterns, engagement levels, and performance data, the system can recommend efficient study routines, breaks, and review sessions, reducing cognitive load by promoting effective time management and reducing decision-making efforts.

2.2 Transfer of Learning

At prima facie, it is worthy for the reader to realize the significance of this section; namely how transfer of learning is and must be the end goal of the educational system. But first, let's crack up a reasonable definition of learning transfer. Transfer of learning is the application of previous knowledge and skills encountered in one context to another context. For the sake of precision, it is worthy of notice that learning transfer is different from ordinary learning in that it requires constancy beyond the context where students learned the material, meanwhile the latter requires temporal constancy (learning something in class and applying it later) . It is a matter of spillover effect mediating through seemingly unrelated contexts. There are several distinctions to learning transfer. For the purpose of this paper, I will be accordingly and rather briefly representing three types of distinctions: 1- Positive vs Negative transfer; Positive transfer is when knowledge in one context enhances performance in other context, negative transfer respectively hinders such a performance. 2- Far vs Near transfer; near transfer occurs when the two contexts are highly correlated or similar, far is the otherwise. 3- High road vs Low road transfer; high road transfer requires abstracting from previous experience concomitant with initial search for connections in novel problems and contexts, low road transfer is stimulus-driven ie; it hinges on galvanizing well-cemented thinking routines by presenting cues(stimulus) that were initially present when the learner formed those routines. From the perspective of educational policy research, positive transfer is the nexus of concern in today's schools because negative transfer can be compensated for with experience (Perkins and Salomon, 1992)[34]. Meanwhile for the far-near distinction debate, research majoritively tilts towards the imbalance of prospect where near transfer is more likely to occur than far transfer (Thorndike, 1923; Scribner and Cole, 1981; Pea and Kurland, 1984)[35] with few exceptions (Clements and Gullo, 1984)[36]. In the wake of the importance of test-taking in not only ensuring that the materials presented in lecture hall is mastered - which they often don't live up to for many reasons discussed in the test-taking section- but in their rather often unfortunately stunted potential in thoroughly preparing, encouraging and evaluating students' abilities to deal with seemingly or literally novel concepts and problems on the fly that can be later encountered in contexts outside of school or university and the role of the teacher must facilitate that transition from fixating on tautologies to creatively fishing out generalities and useful abstractions from the test (Lindquist, 1951)[37], it is essential therefore to shed light on the conditions under which transfer occurs; 1- Deliberate repeated practice; it has been shown that deliberate practice is what distinguishes between experts and non-experts (Tetlock, 2005)[38]. That makes perfect sense, if one practices areas where one feels uncomfortable the most, positive far transfer occurs. 2- Explicit abstraction; this is the idea of thoroughly abstracting from past experiences in regard to usefulness ie; asking the question "What is it that I can use from solving this problem or learning this concept that can help me with solving another problem or learning another concept?" . 3- Active self-monitoring, which is also called Metacognition, which is the awareness of one's own learning and understanding of a concept (Flavell, 1979). 4- Arousing Mindfulness 5- Using metaphors/analogies.

2.3 Learning styles

One of the greatest myths in education that stood the test of time is that of learning styles. There are two naive fundamental assumptions operating under this theory; first that each student has their own learning style that remains constant across different subjects, second that student's learning performance will increase if the instructor presents information in the mode that speaks to their learning style. There are several version of learning style theory, the most prominent one is perhaps that of the VARK model (Visual, Auditory, Reading and Kinesthetic) (Fleming and Braume 2006)[39]. In the words of Neil Fleming himself "During this time I watched some 9000 classes. I was puzzled when I observed excellent teachers who did not reach some learners, and poor teachers who did....There are, of course, many reasons for what I observed. But one topic that seemed to hold some magic, some explanatory power, was preferred modes of learning, 'modal preferences' ". Unsurprisingly, evidence for learning styles (sometimes called Meshing Hypothesis) is lacking (Willingham, 2015; Rogowsky, 2015)[40]. Nonetheless, students and even worse instructors and teachers buy into it because it assumes students to be different - which is obviously true, except for other reasons- but that it also has this egalitarian atmosphere which perpetuates its widespread proliferation making it a "common knowledge", 90 percent of teachers from several countries believe in learning style theory (Willingham and Dobolyi, 2017)[41]. However, and from what we already know, the difference between students may not be due to some innate learning preferences but of ability, background knowledge, interest as well as the content of the topic in question. In the wake of all that, the focus should be rather directed on empirically supported parameters, that is to say that teachers/instructors shall emphasize the difference in abilities (equivalently the time required to master a subject), background knowledge and interests between students in a mastery-based framework that prioritizes the needs of students (Bloom, 1968) instead of assessing educational performance based on bogus concepts that are not evidence-based nor applied properly which may probably end up shoehorning students into categorical buckets thinking they will only learn in a strictly, or rather vaguely for that matter, defined mode of information. Subsequently on a practical instructional level, the use of technology (like digital media) for example or multimedia learning shall accommodate students' abilities, interests and background knowledge.

2.4 Learning proxy variables

Another equivalently mythical but surreptitiously bamboozling held belief among educators, instructors and policymakers comes from the irresistible and rather erroneous tendency of hollowing learning out of its invisible and unconscious nature via the wishful thinking of making it "observable". Just like the meshing hypothesis, learning proxy variables (engagement, motivation, interest, business...etc) miss an awful a lot of insights to whether learning occurs or not (Coe, 2017)[42]. This is rather obvious because the very idea of a proxy variable predicates on a correlation not a causation. Even worse, if one has two variables that highly correlate but one variable is extremely more difficult to understand than the other, then the correlation becomes an oversimplification of the difficult-to-understand parameter in question (learning in this case) thinking that we can get away with it as a consequence. Now and to be fair to these proxy variables, it's worth stating that learning requires these conditions to occur but such conditions per se don't necessarily signify a process of learning going on. To recap, proxy-based approach to instructional teaching is one of the persisting gaps in pedagogy today; foolishly relying on factors for the sole virtue of being easily and directly measurable, observable and - one could argue- objective is naive optimism bias stemming from an outdated Pavlovian wishful thinking. As the economist Milton Friedman puts it: "There ain't no such a thing as a free lunch"

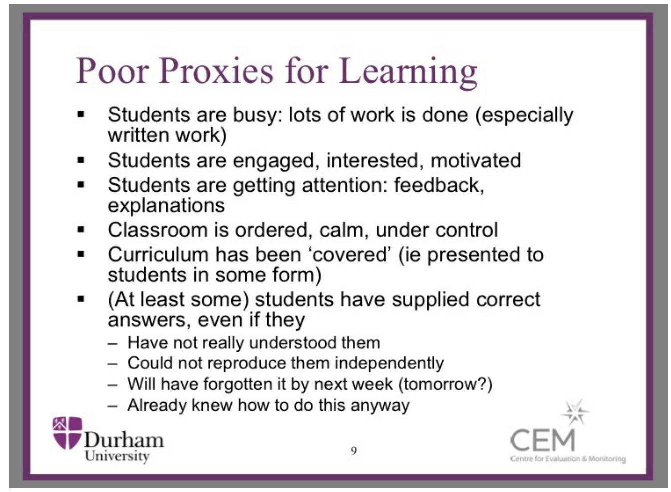

Figure 7. Learning proxy variables

2.5 Learning theories

There are a plethora of learning theories in the literature. In the scope of this paper, I am going to be presenting only three pertinent learning theories that are of essential importance to the research question solely in the light of their applications in education; namely Behaviorism, Cognitivism and Connectivism. . Behaviorism is a psychological approach that focuses on observable behavior as the primary subject of study, disregarding internal mental processes. Key figures in behaviorism include B.F. Skinner, Ivan Pavlov, Edward Thorndike, and John B. Watson. Classical conditioning, proposed by Pavlov, involves the learning of associations between stimuli and responses. It suggests that neutral stimuli can elicit reflexive responses when paired repeatedly with an unconditioned stimulus that naturally triggers the response. For example, Pavlov's famous experiment with dogs demonstrated that dogs could learn to associate the sound of a bell (neutral stimulus) with the presentation of food (unconditioned stimulus), leading to salivation (unconditioned response) upon hearing the bell alone (now conditioned stimulus). Operant conditioning, developed by Skinner, focuses on the consequences of behavior in shaping future behavior. It emphasizes the role of reinforcement (rewards or punishments) in strengthening or weakening the likelihood of a particular response occurring again. Skinner's work introduced the concept of operant conditioning through his experiments with rats and pigeons in a Skinner box, demonstrating how behaviors are shaped through reinforcement schedules, such as positive reinforcement (rewarding desired behavior), negative reinforcement (removal of an aversive stimulus), punishment (application of aversive stimuli), and extinction (the fading away of a behavior when reinforcement is removed). Thorndike's law of effect posits that behaviors followed by satisfying consequences are more likely to be repeated, while those followed by unsatisfying consequences are less likely to recur. Lastly, John B. Watson is known for advocating the belief that behavior is primarily a result of conditioning and environmental factors, discounting the influence of innate factors. His research aimed to demonstrate how behaviors could be learned and manipulated, highlighting the significance of environmental stimuli in shaping behavior.

Behaviorism has wide applications in education. It predicates on the idea that learning occurs when there is an observable change in behavior resulting from conditioning students to stimuli. Taken at surface value, this might be useful in lecture halls or TD sessions as to rewarding students for exhibiting desirable behaviors like active participation or constant concentration, though that has its drawbacks ranging from conflating learning with those "desirable behaviors" to overlooking equally significant factors like background prior knowledge and innate ability, but most chiefly it can be used in test assessment for example. Behaviorism is not as prominent today as it was decades ago and that's for a good reason, namely the fact that it needs to take in consideration radically important aspects of learning that it once overlooked as to allow for an eclectic mixture with other theories as a practically additional leg-up but that at the same time conserves its unique perspective on how students learn and more crucially that it shall be applied properly (Skinner, 1915; Kahneman and Tversky, 1979)[43]. However, and for the purpose of this paper as well as the educational convictions of the author and the educational state-of-the-art, Behaviorism remains outdated as a model to be used to thoroughly explain how learning relates or not to specific educational variables outlined in this work. That's because Behaviorism is hinging on a blackbox outlook on learning and human subject at large. However, that doesn't rule out its practical usefulness in today's classrooms or lecture halls, as George Box puts it : "All models are wrong, but some are useful". Moreover, from a pedagogical point view, Behaviorism is a passive form of teaching where the teacher is the "sage on the stage" dictating all aspects of the classroom leaving no chance for students to think for themselves (P. Freire, 1964 "Pedagogy of the oppressed")[44]. There are reasons why the Skinner learner machine failed so miserably to meet student's demands prior to the electronic Computer Aided Instruction (CAI) revolution. Connectivism (G. Siemens, 2004)[45] is a theoretical framework of learning in the digital age (e-learning) that tries to examine how the Web 2.0 technologies (forums, web browsers, search engines, wikis...etc) helped foster learning in a high scalable and global scope setting by providing knowledge and learning available in the form of external databases or knowledge graphs as to make connections between all these different sets of information (favoring connections over current knowledge state) ie; knowledge as an externalized network or a graph database and learning as a process of pattern recognition of all the connections/nodes of knowledge, and using it to perform an action. Obiouvlsy, the stronger and the more numerous the connections are, the better the learning (Metcalfe's law). In the context of this paper, connectivism as a framework is even more empowered thanks to the emerging Semantic Web (Web 3.0) architecture of the World Wide Web thanks to generative AI and AGI, where learning is now extended to a more interactive, automated and personalized layer that enriches the student's learning experience in a way that's akin to one-to-one tutoring, something that was not possible within the realm of Web 2.0 technologies. Cognitivism on the other hand pictures learning as an internal mental process and subsequently tend to focus on information distribution from the teacher to the student tilting the attention on what goes on inside students' minds; how information is received, organized, stored or retrieved but most importantly, and for the sake of the research problem, how information is processed and at what level. Unlike Behaviorism, ensuring that information is well transmitted to internal cognition remains at the heart of assessing the quality of education. But what is processing? There is no strictly defined metric about what information processing is, but a plausible definition would be unpacking, deconstructing information into small bits and reverse-engineer from there. That said, let's draw the distinction between shallow processing, which is unfortunately very common in today's classrooms, and deep processing. But before that, it's worth noting that this distinction is radically crucial to understand for this research problem, though not the only key concept, but it explains an awful a lot of why some students skim the cream while others struggle to no end, there is no magic spell here. Anyway, there are three types of processing levels but that's fairly far-fetched. In fact, processing levels distributed on a continuum spectrum; where one is on that spectrum depends by and large on the orienting task in question (Nolen, 1988)[46] and attention span as well as memory recall (Craik and Lockhart, 1972; Craik and Tulving, 1975)[47] though some suggests that elaboration level is responsible for such a spurious association (Boyd, 1986)[48]. Shallow processing (also called structural) is unpacking superficial aspects of information which are the sensory perceptual features purported to incite the senses only. Intermediate processing (Phonemic) is the phase of actually recognizing and labeling the information in question. Deep processing (Semantic) capitalizes upon the underlying logic, meaning and patterns behind the information in question, this works through asking a plethora of pertinent, meaning-specific questions in an Ishikawa diagram kind of approach; What is it saying? Why is true? When is it useful? How does it relate to other concepts? How is it different from other concepts? How does it relate to personal experience or real life? What are the kind of information modes that could represent it (visuals, diagrams..etc)? What are all the levels of abstractions to present it? Where learning about it can help in other fields?..etc. Bluntly put, deep processing is the systemization, mechanization and algorithmization of what it means to be intelligent or "genius". The genie is back to the bottle as it's clear now what the difference between processing levels amount to, practically or otherwise. The discussion of this paper largely builds on this concept.

Not far from Cognitivism – and in fact a branch of it-

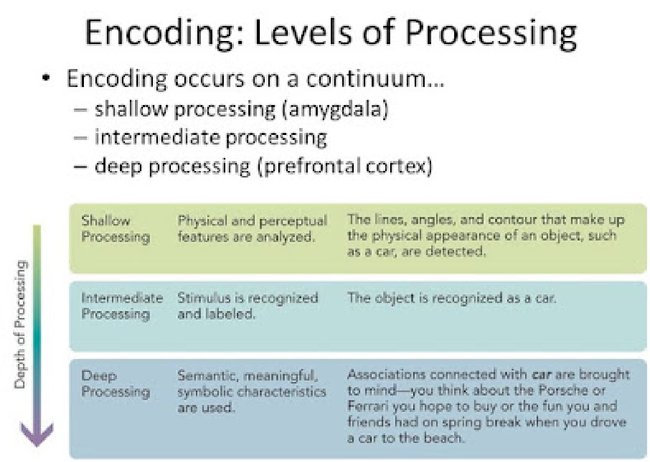

Figure 8. Processing levels

is Constructivism. Initially pioneered by Jean Piaget, Constructivism is the school of thought that places background knowledge, side-to-side with ability, at the forefront of explaining why students differ in educational performance, learning difficulty and their overall interaction with school at large. Students learn by "constructing" new meaning or "schemata" based on prior knowledge, an intuitive conclusion following that is that knowledge is messily individualized, idiosyncratic, subjective and stochastically unsystematic. Essential to this school of thought is the notion of Bloom Taxonomy (Bloom, 1956)[49] which is the idea that learning comes in waves or stages with different thinking orders; low order thinking and high order thinking. Low order thinking include aspects of learning like remembering, understanding and applying prior knowledge to new situations whilst high order thinking involves contrasting information and connecting it to other concepts, evaluating and creating. Though there are other revised versions of Bloom Taxonomy (Anderson and Karathwohl, 2001; Chuches, 2008)[50] and even similar concepts like SOLO taxonomy (K. Collis and J. Biggs, 1982)[51], the essence is still the same. Bloom Taxonomy is a concept to educate by, it reframes the educational endeavor in that the goal of the students is to climb the Bloom pyramid with teachers facilitating the process, by taking into account the shift in attention and responsibility in learning levels between highschool and college (McGuire, 2015)[52]. In the scope of this paper, Bloom Taxonomy is used to operationalize, concretize and discretize the learning experience of students into different measurable, quantifiable categories that we know and understand to a great extent.

Akin to Bloom taxonomy is the SOLO (Structure of the Observed Learning Outcome) taxonomy which is also worthy of mention though it has relatively less significance on the level of solution component, SOLO is a framework used to classify and evaluate the complexity of learning outcomes. It consists of five levels: pre-structural, uni-structural, multi-structural, relational, and extended abstract. The pre-structural level represents a lack of understanding, while the uni-structural level

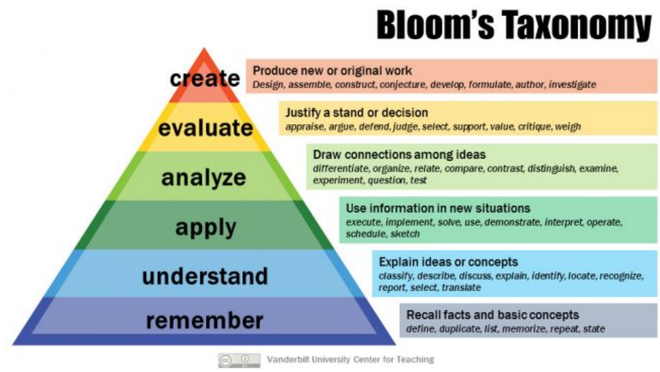

Figure 9. Bloom Taxonomy (Bloom, 1956)

indicates the understanding of only one aspect of a concept. At the multi-structural level, the learner can comprehend multiple aspects but fails to connect them. The relational level reflects the ability to understand the connections between different concepts. Finally, the extended abstract level represents the highest level of understanding, where the learner can apply the knowledge in novel and abstract ways. The SOLO taxonomy helps educators assess and scaffold learning progress, guiding students towards deeper levels of understanding.

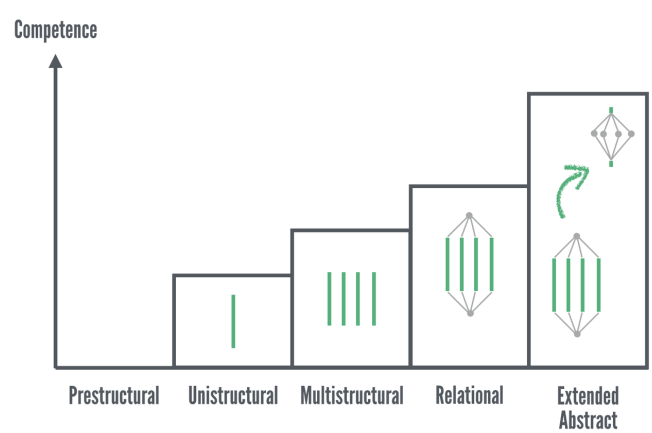

Figure 10. SOLO taxonomy (K. Collis and J. Biggs, 1982)

3. Research Problem

A well-constructed student model is fundamental for designing effective educational interventions. Without an accurate understanding of a student's learning needs, gaps in knowledge, and cognitive strengths, interventions risk being misaligned, which leaves these gaps unaddressed. Over time, this accumulation of unmet learning needs can lead to frustration, disengagement, and burnout. Therefore, identifying and representing the student's learning profile with precision is critical for ensuring interventions are targeted and meaningful, ultimately supporting long-term learning outcomes. The core challenge lies in the lack of a robust and reliable student model, which prevents the timely and effective support of students. Here is a graph outlining the problem in its core logic:

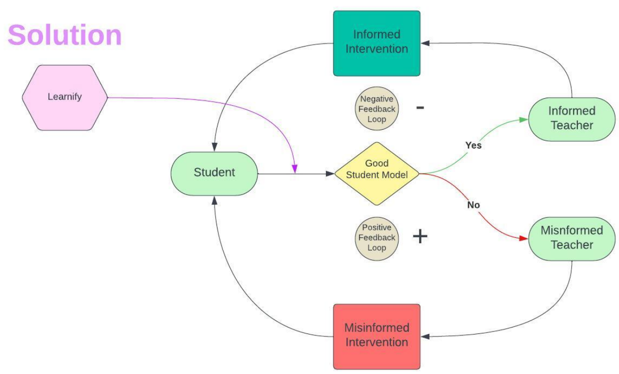

Figure 11. Visual Representation of the Problem

3.1 *Pedagogical Problem: Bloom 2 sigma problem*

The Bloom's 2 sigma problem refers to a challenge identified by educational psychologist Benjamin Bloom. In a seminal study published in 1984, Bloom compared the learning outcomes of students who received one-on-one tutoring (referred to as the "2 sigma" condition) with those who received conventional classroom instruction in one-size-fits-all lecturing model. The study found that students who received personalized tutoring performed, on average, two standard deviations (2 sigmas) above students in the traditional classroom setting.

Results have been documented in the famous Bloom's 2 sigma problem (Bloom, 1984)[54] suggesting that the average student tutored one-to-one using mastery learning techniques (Bloom, 1968) performed two standard deviations better than students educated in a classroom environment with one teacher to 30 students, with or without mastery learning (more precisely, the average tutored student was above 98 percent of the students in the control class. Moreover, about 90 percent of the tutored students ... attained the level of summative achievement reached by only the highest 20 percent" of the control class). Briefly, Bloom's 2 sigma problem focus on two fudamental parameters or variables: student size (thus teacher/individual student attention ratio via comparing individual attention to students versus group performance assessment, also non-individualized learning pace for each student as well as averaging out different learning rates by methylating unsystematic factors resulting from non-shared environmental factors (Plomin and Daniels ‘‘Gloomy Prospect”, 1987) –like interest, background knowledge, prerequisite level, ability, which goes against the principles of Mastery Learning (Bloom, 1968) that sets time required for student's learning curve of achieving the same level of mastery as its most important parameter) and instructional method

In the light of the famous Bloom's 2 sigma problem, B. Bloom concluded that one-to-one tutoring can't be scaled up without accumulating cost due to the fact that it has a non-zero marginal cost. Therefore, the sort of approach he proposed is a complex coordination of alterable variables (Learner, Instructional Material, Home Environment/Peer-Group and Teacher) hinging on the hopes of finding a group instruction method that results in the same performance as one-to-one tutoring through finding the just-right combination of these alterable variables. The following is a table summarizing the results found by B. Bloom and his graduate students.

Bloom's 2 sigma problem is amounted to a sort of "educational disease" as it obeys the following criteria: 1- It is easy to state and understand. 2- It seems to be accessible for resolution via mere lateral thinking (re-shaping of existing approaches and methods in the just-right way). 3- It has been repeatedly "solved" by startups, companies and organizations in the field of education (examples include TutorOcean, KhanAcademy, CarnegieLearning...etc)

In the light of the famous Bloom's 2 sigma problem, B. Bloom concluded that one-to-one tutoring can't be scaled up without accumulating cost due to the fact that it has a non-zero marginal cost. Therefore, the sort of approach he proposed is a complex coordination of alterable variables (Learner, Instructional Material, Home Environment/Peer-Group and Teacher) hinging on the hopes of finding a group instruction method that results in the same performance as one-to-one tutoring through finding the just-right combination of these alterable variables. The following is a table summarizing the results found by B. Bloom and his graduate students:

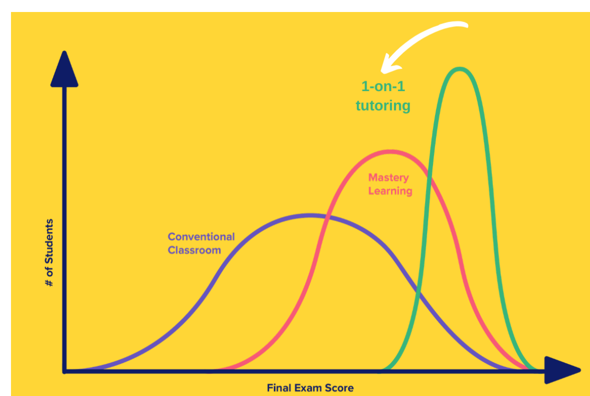

Figure 12. Achievement distribution for students under conventional, mastery learning and tutorial instruction

Object of change process	Alterable variable	Effect Size	Percentile equivalent
Teacher	Tutoring instruction	2.00	98
Teacher	Reinforcement	1.2	
Learner	Feedback-corrective (mastery learning)	1.00	84
Teacher	Cues and explanations	1.00	
Teacher, Student	Student classroom participation	1.00	
Learner	Student time on task	1.00	
Learner	Improved reading/study skills	1.00	
Home environment / peer group	Cooperative learning	0.80	79
Teacher	Homework (graded)	0.80	
Learner	Initial cognitive prerequisites	0.60	
Home environment / peer group	Home environment intervention	0.50	69

Figure 13. Effect of selected alterable variables on student achievement

3.2 Technology Problem: Domain Expertise and LLMs

The use of large language models (LLMs) to personalize learning holds considerable potential, but it also brings with it significant limitations, particularly when LLMs are tasked with domain expertise. One of the major challenges in using LLMs for this purpose is their tendency to hallucinate or generate incorrect information. While LLMs can provide coherent and contextually relevant responses, they do not possess true understanding or reasoning capabilities. This lack of comprehension means that LLMs can produce misleading or inaccurate answers, especially in more specialized or complex domains. In the context of personalized learning, this can be especially harmful, as students may inadvertently rely on incorrect information, which can distort their understanding and ultimately hinder their learning progress. The ability of LLMs to produce faulty content, without a mechanism to recognize or correct errors, makes them unreliable when serving as domain experts.

Furthermore, a more subtle but equally critical issue arises when LLMs are positioned as the experts in a specific domain. This can unintentionally undermine student autonomy. Traditionally, in personalized learning, the role of a tutor is not merely to deliver information but to foster critical thinking and guide students toward self-directed learning. However, when LLMs take on the role of a domain expert, they risk providing direct solutions to problems instead of encouraging students to engage with the material themselves. This can result in students passively receiving information, rather than actively constructing their own understanding. By outsourcing problem-solving to the model, students may become reliant on the LLM for answers, which reduces their ability to think critically and solve problems independently. This approach conflicts with the educational goal of developing autonomy in learners, as it substitutes the student's cognitive efforts with the model's output, diminishing the student's opportunity to grow and develop essential skills.

To address these concerns, we propose a shift in the use of LLMs within personalized learning environments. Rather than relying on the LLM to be an expert in the domain, we suggest positioning it as an expert on the **student model**. In this framework, the LLM's role is not to provide direct domain knowledge, but to understand the learner's strengths, weaknesses, and cognitive profile, then guide the student in navigating these challenges. The LLM would act as a **meta-tutor**, offering metacognitive guidance, helping students identify their learning gaps, and providing strategies for overcoming them. This would allow the LLM to help the student understand how to approach their learning and adapt strategies according to their individual profile, rather than simply presenting domain-specific content.

By shifting the focus of the LLM from being a domain expert to a facilitator of metacognitive growth, the model can help students better understand their learning process. For example, if a student struggles with recalling information, the LLM could recommend strategies like spaced repetition or retrieval practice. If the student faces difficulty in analyzing complex concepts, the LLM might suggest methods for breaking down problems into smaller, manageable parts. This approach helps students become more self-aware and self-regulated in their learning, focusing on how to learn rather than just what to learn. Crucially, this model preserves the student's autonomy, as the LLM guides the student to take charge of their learning process, rather than simply providing them with answers.

Ultimately, this solution ensures that the LLM's involvement in the learning process is one that promotes independence and critical thinking, rather than replacing the student's cognitive efforts with its own output. By leveraging the LLM's ability to understand and support metacognitive strategies tailored to the student's learning profile, we can create a system that empowers the learner while mitigating the risks of hallucination and passive learning. This model of using LLMs as meta-tutors instead of domain experts represents a significant shift toward ensuring student autonomy, promoting deeper learning, and offering a more sustainable and effective approach to personalized education.

4. Methodology

As stated previously, this paper is primarily concerned with how seemingly irrelevant, accumulating gaps affect the learning experience of students which are the unit of analysis of this research. First we discuss selection criteria used to choose what specific subproblems to work on within each educational variable. Second, an empathy map along with research questions/hypotheses are presented. Afterwards, I discuss the research design, data collection and analysis.

4.1 Selection Criteria for Educational variables and their subproblems:

1- Do-able within the budget, timeframe and principle (there is a room for improvement) -> Excluding factors like natural ability, school infrastructure..etc
2- Student's experience with educational variables must be systematically coercive and non-contingent, non-sporadic and casual.
3- The subproblems within each educational variable must be within the academic reach of only Cognitivism or Constructivism in terms of modeling, but there is no strict selectivity in terms of problem-solving though it's preferable to be framed in Cognitivist/Constructivist language or eclectically fused with other approach (Behaviorism is a great example here) -> Excluding factors like interactivity, engagement (emotional and behavioral), classroom/lecture hall level of calmness and order 4- The subproblems within each educational variable mustn't have learning styles theory as its basis.
5- The subproblems with each educational variable must have an empirically supported substantive, causative, definitional or intrinsic association and relationship with learning and not just proxy variables or correlations. -> Excluding factors like motivation (extrinsic and intrinsic), engagement, classroom/lecture hall level of calmness and order and almost all learning proxy variables discussed above (Coe, 2017)
6- The subproblems within each educational variable must be

conceptually distinct from the operationalization of the learning experience (Kahneman theory, Transfer theory, Multiple store model of memory, Cognitivist theory, Bloom Taxonomy, SOLO taxonomy) as to avoid tautology bias.
7- The subproblems within each educational variable must have the functional state of an input, a cause or at least an independent predictor variable. This criterion is not mandatory as this research paper is attempting to draw correlations between variables regardless of causality arrow, but it's preferable that this condition is checked.
8- The subproblems presented within each educational variables must form a logistic management problem, i.e, an optimization problem of ensuring the flow of goods/services (information) from the point of origin (teacher, instructor or any source of information) to the point of consumption (by students).

4.2 Student's learning experience : Empathy Map/Ishikawa Diagram

Figure 14. Empathy Map representing the learning experience of students

4.3 Research Questions and Hypotheses

1- Is there a relationship between information distribution during lecture and learning acquisition? 2- Is there a relationship between teacher's pedagogy during lecture and learning acquisition? 3- Is there a relationship between operational variables during practice problems session and learning acquisition? 4- Is there a relationship between instructor's awareness during practice problems and learning acquisition? 5- Do teaching to the test affect the learning process? 6- Does the nature of test preparation affect the learning experience?

To answer that, a thoroughly appropriate methodology must be followed so as to unravel the holes of the educational swiss cheese.

4.4 Operationalization

Operationalizing Bloom's Taxonomy involves transforming the theoretical framework into a practical tool for generating quantitative survey questions. Bloom's Taxonomy, a widely used educational framework, classifies learning objectives into different cognitive levels, such as knowledge, comprehension, application, analysis, synthesis, and evaluation. To operationalize Bloom's Taxonomy for survey generation, one can utilize its cognitive levels as a guideline to design questions that assess different levels of understanding and thinking. The operationalization process begins by mapping each cognitive level to specific question types or prompts that align with the desired learning outcomes. For example, knowledge-level questions may focus on factual recall, comprehension-level questions may assess understanding and interpretation, while application-level questions may require applying knowledge to solve problems or make connections. Furthermore, operationalization can involve incorporating appropriate response scales or formats to capture respondents' levels of proficiency or agreement. This may include Likert scales, multiple-choice options, or ranking exercises that reflect the intended cognitive level being evaluated (in our context, a Likert scale was chosen for convenience reasons).

By operationalizing Bloom's Taxonomy in survey design, it is possible to gather quantitative data that enables to assess and compare learners' knowledge, skills, and abilities across different cognitive domains. This approach facilitates the measurement of learning outcomes, identifies areas for improvement, and provides valuable insights for instructional design and evaluation. It helps generate targeted and meaningful survey instruments that effectively capture learners' cognitive development and progress.

4.5 Research Design

The methodological approach used is chiefly aimed at establishing a correlational relationship between students' systematically coercive and non-contingent, non-sporadic and casual experiences in each learning bottleneck discussed previously and learning difficulty, an average is obtained as to account for the overall association between learning acquisition and the aforementioned educational variables. Though previous literature provides empirical evidence of the far-reaching influences of each of those bottlenecks on their own, this paper offers new primarily collected quantitative data to these old problems but in an aggregated setting ie; constructing a single statistical model encompassing all parameters combined which allow to infer conclusions about the model-independent effects of each of those factors as to compare between them, to average out inter-correlations and compute intra-correlations (between the bottlenecks themselves). This correlational research method is one of the three prominent methods in educational research along with experimental and descriptive because it's a standard methodology that is widely used specifically in identifying latent educational variables.

4.6 Data collection

First, it is important to mention the study unit of data collection: Students. The participants are students from engineering schools.
The survey consisted of 3 multiple-choice questions and questions measured on a 5-point Likert scale, the survey doesn't include Yes/No questions given the informational limitations resulting from that. Given the huge amount of questions in total, a divide and conquer approach was used in online distribution.This survey is purely student/learner-centered as it tends to focus exclusively on the experience of students with the targeted educational variables in this paper. The survey was conducted online in an anonymous and self-administered setting as to reduce social desirability bias which may hinder the validity and reliability of the research. Also, the chances of undercoverage bias are sufficiently reduced thanks to the fact that most students targeted use the internet and could easily answer the survey. Moreover, the purpose of the survey was hidden from the participants as to avoid demand characteristics bias that may alter their attitudes towards the questionnaire (Hawthorne effect). The participants filled the survey at the comfort of their pace without intervention as to avoid any pressure that can result from interviewer bias, but also that the questions were clear, close-ended allowing no room for misunderstanding, neutrally worded. In the light of the abstract nature of "learning difficulty", It's been decided to operationalize it through the criteria/discriminant validity notion by asking questions about concrete concepts that commonly correlate, negatively or positively, with learning quality (Bloom Taxonomy is a framework for operationalization)

The goal was to virtually collect survey responses from students attending engineering schools or universities regardless of their majors. The sampling method used is random sampling with the only criterion required is that the participants are students of a university or a school. Thanks to the principle of maximum entropy, random selection allow for establishing sturdy statistical correlations devoid of sampling bias. On another footnote, the data obtained is triangulated ie; coming from different sources (3 different schools) so as to minimize the chances of coincidental or biased answers as well as controlling for personal characteristics (ability, background knowledge, interests..etc) and eliminating spuriousness (law of large numbers). In that telling, it is a good practice to categorize the questions into two groups: Student learning experience questions and school variables questions. In that telling, here are the questions:

Student's Learning Experience:

1-You get to effectively understand what you learned
2-You get to apply what you learned in new situations
3-You get to connect the dots or find similarities between seemingly irrelevant ideas, courses, fields
4-You feel that you lack prerequisites for the class
5-Answering your test questions relies on passive/rote learning (remembering answers to similar questions or results from previous exams)

School Variables:

1-You often feel rushed during the lecture due to high rate of information distribution
2-You feel that the amount of information during the lecture is often too large
3-Your class uses multimedia sources of information (textbooks/books, lecture videos, MOOCs, online tutorials, encyclopedias)
4-Your lecturer uses clickers/quizzes before, during, or after lecture with feedback
5-Your lecturer gives a quick overview (wrap-up) of the previous lecture before starting
6-Your lecturer asks you to try to learn the course before the lecture
7-Your lecturer helps fill the prerequisite gaps if you have any
8-You feel your instructors teach to the test (i.e., tests overshadow the learning experience)
9-Productive time covers up the total time of practice problems session (i.e., the effort input you exert in the practice problems session is efficiently converted into a desired product)
10-You feel that the practice problems session doesn't fit your requirements (either underfitting or overfitting)
11-Your instructor assumes everyone already mastered their courses during practice problems

Here, a student is defined as a person who is involved in one or some combination of the areas outlined in the research question. Participants were given indefinite duration to fill in the survey anonymously. In total, students responded, but not all surveys were fully completed. Due to this, survey results were included in the analysis. But without further ado, here are the results of the data collected:

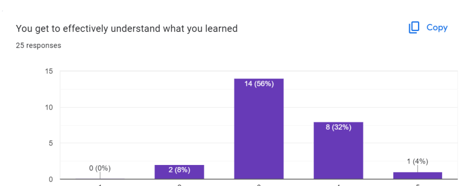

Figure 15. Comprehension data

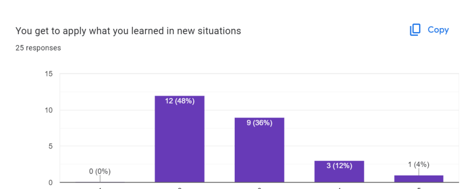

Figure 16. Application data

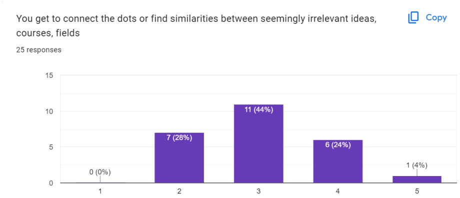

Figure 17. Analysis data

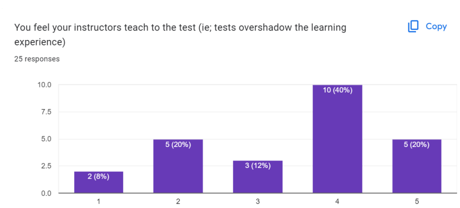

Figure 18. Test Surrogation effect data

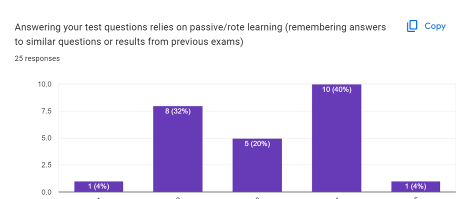

Figure 19. Test validity data

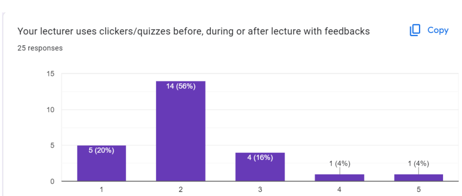

Figure 20. Clickers Feedback data

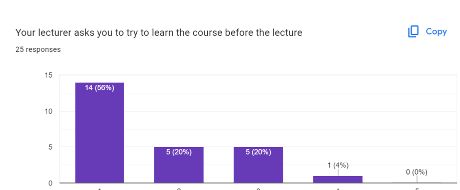

Figure 21. Flipped Learning data

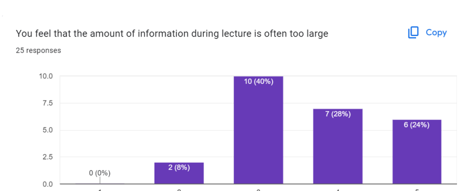

Figure 22. Information Size during lecture data

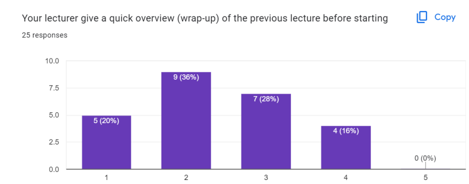

Figure 23. Quick Overview data

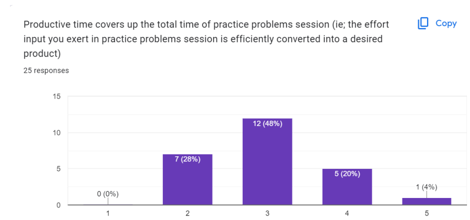

Figure 24. TD efficiency data

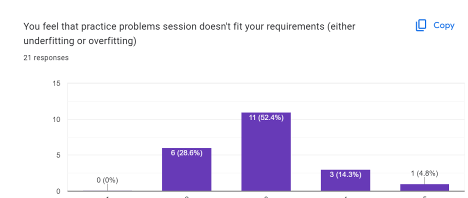

Figure 25. Requirements Misalignment data

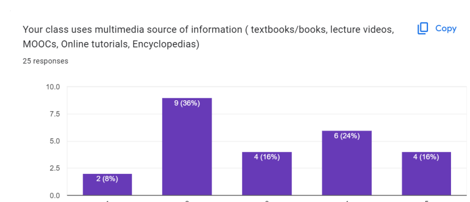

Figure 26. Multimedia learning data

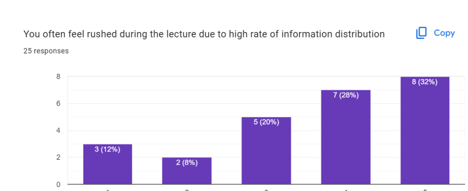

Figure 27. Lecture Rush data

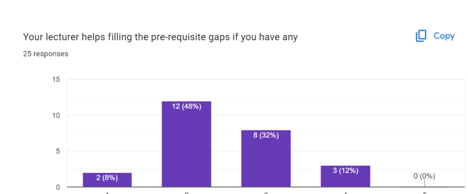

Figure 28. Learning Gaps filling data

Figure 29. Lacking Prerequisites data

Figure 30. TD instructor assumption data

4.7 Data analysis

Data was prepared using SPSS software. Missing values were checked for and removed. After representing the boxplot of the data, few outliers were displayed and removed eventually as to eke out for statistical bias. The statistical tests used are Normality tests and Correlation test: First, the Normality test (Shapiro–Wilk test and Kolmogorov–Smirnov test) and Spearman's correlation test (non–parametric test). For the correlation tests, a two–tailed test was used in terms of statistical significance of the test statistic (bidirectional). The justification for using the Spearman's correlation test is the fact that the data collected is of ordinal nature (Likert scale) which is compatible with Spearman's correlation test as opposed to say, Pearson's correlation test or Kendall's tau correlation.

4.8 Justifying and evaluating the research design

One of the limitations of this methodology that are worth mentioning is the use of correlations to infer conclusions. As commonly known, correlations don't necessarily entail causation "cum hoc ergo propter hoc" which makes it harder to identify the exact processes that are responsible for the results. Another caveat for using a correlational approach is the possibility of spurious relationships. Moreover, it is noteworthy that using surveys for data collection is intrinsically prone to question (divide and conquer) order bias which sometimes ends up shepherding participants to tilt to one side of the answer. However, and in order to eliminate such a bias, there was automatic shuffling of questions from different areas. But perhaps the most hazardous bias plausibly inherent in this research is tautology bias which suggests that perhaps the correlation between the educational variables and learning difficulty is merely due to definitional conflation between the two resulting from each participant subjective interpretation of the question which is very common in questionnaires validity problem; to give but one example, the participants might've interpreted their grapple with say for example difficulty receiving information during lecture (an educational variable) as a learning difficulty itself. In other words, learning difficulty is not necessarily some educational variables independent abstract notion with its own measures. Nonetheless, the methodological approach used provides an unprecedented chance of incorporating the aforementioned educational variables in one single model so as to allow for a meaningful comparison in their contribution to the learning experience while also obtaining the overall effect of such variables on the learning experience with PCA approach as well as how much they interact between each others organically.

5. Results

The results indicate that there is a correlation between the school variables and students learning experience. However, due to data shortage (which is going to be solved later via data augmentation and retesting by focusing solely on the Bloom Taxonomy variables), the correlation is not statistically significant (.915)

Nonparametric Correlations

Correlations

			Student_Learning_Experience	School_Variables
Spearman's rho	Student_Learning_Experience	Correlation Coefficient	1.000	-.030
		Sig. (2-tailed)	.	.915
		N	25	15
	School_Variables	Correlation Coefficient	-.030	1.000
		Sig. (2-tailed)	.915	.
		N	15	15

Figure 31. Spearman's correlation test between Students Learning Experience and School Variables

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Understanding	.258	15	.008	.881	15	.049
Application	.366	15	.000	.705	15	.000
Analysis	.238	15	.022	.817	15	.006
Prerequisites_Gaps	.237	15	.023	.881	15	.049
LectureRush	.276	15	.003	.816	15	.006
ClickersUse	.314	15	.000	.842	15	.014
InformationSize	.244	15	.017	.866	15	.030
FlippedLearning	.361	15	.000	.731	15	.001
Prerequisites_Gaps_Filling	.316	15	.000	.790	15	.003
MultimediaLearning	.280	15	.003	.862	15	.026
Passive_Testing	.256	15	.009	.882	15	.051
TD_Efficiency	.212	15	.068	.817	15	.006
Teaching_To_Test	.341	15	.000	.727	15	.000
Instructor_Mastery_Assumption	.314	15	.000	.842	15	.014
QuickOverview	.306	15	.001	.846	15	.015
TD_Requirements_Misalignment	.242	15	.018	.828	15	.009

Figure 32. Normality test: Kolmogorov-Smirnov test and Shapiro-Wilk test

Regardless of the correlation significance level, another equally important measure is the skewness of the Bloom Taxonomy data distribution (to be discussed in the Discussion chapter)

6. Discussion

In our proposed solution, we emphasize the critical need for a personalized student model that transcends traditional domain-specific approaches in Intelligent Tutoring Systems (ITS). By focusing on the student's core learning traits rather than isolated academic domains, we offer a more holistic and adaptive framework. Our model is grounded in Bloom's Taxonomy, leveraging its tiered structure to scaffold learning experiences. This approach stands in contrast to other models, like BloomBERT and Khanmigo, which limit themselves to specific domains, thus ignoring the broader, foundational aspects of learning.

Our reliance on large language models (LLMs) for student modeling, while leaving domain expertise to students and teachers, creates a balance between personalization and autonomy. This divergence from the norm ensures safer, more accurate models without the risks of LLM hallucinations in specialized contexts. The solution allows students to explore their learning journey in a more self-directed manner while still benefiting from adaptive scaffolding, reducing the dependency on rigid, predefined learning paths.

Furthermore, our solution is designed to scale globally, focusing on foundational components of education such as the student model. This allows for broader applicability, addressing systemic challenges like those seen in Morocco's low PISA ranking, which underscores the importance of targeting educational frameworks rather than specific content domains. The solution's scalability is rooted in its ability to cater to diverse student profiles across regions and educational systems, ensuring that it can address educational disparities on a global scale.

In conclusion, our approach offers a more sustainable, adaptive, and holistic path forward for ITS development. By shifting the focus from domain knowledge to learning itself, we propose a model that adapts to the learner's needs, encourages autonomy, and provides a robust framework for scalable education technology. This paradigm shift has the potential to significantly improve the learning experience, particularly in underserved educational contexts, while contributing to the broader field of EdTech innovation.

7. Solution

7.1 Conceptual Representation

Our approach's impact extends beyond immediate personalization. By abstracting learning into a universal metric ("Bloom Units") and creating a scalable, student-centric model, the system is adaptable across domains, educational levels, and cultural contexts. Scalability is inherent in its reliance on information—naturally diffusive, anti-rival, and resource-efficient—which can be shared across teachers, peers, and students themselves.

With a feasible design rooted in Bloom Taxonomy and a focus on core learning mechanisms, our solution not only addresses systemic educational challenges but also lays the groundwork for a transformative EdTech paradigm shift—one that does more with less by targeting the kernel of learning itself.

8. Conclusion

This work proposes a paradigm shift in Intelligent Tutoring Systems (ITS) by redefining the focus from domain-specific

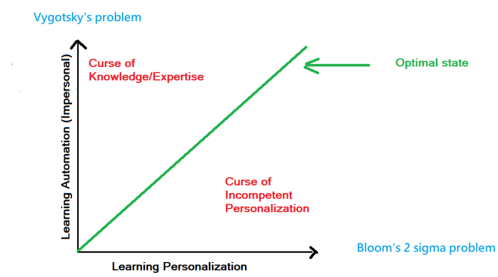

Figure 33. Graph illustrating the relationship between learning effort and personalization

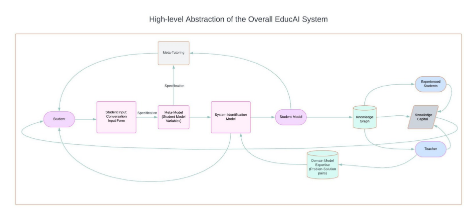

Figure 34. Logical Architecture of the overall product

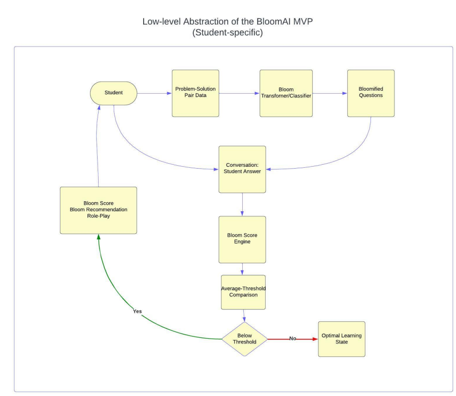

Figure 35. Logical Architecture of the MVP

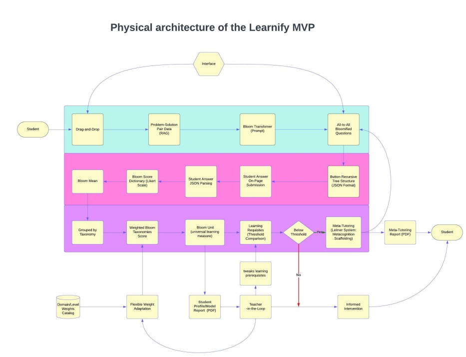

Figure 36. Physical Architecture of the MVP

expertise to a student-centric approach grounded in universal learning principles. By leveraging Bloom Taxonomy to model and personalize student profiles, our solution addresses the core challenges of learning variability and accumulated learning gaps—issues often overlooked in traditional ITS designs.

Our contrarian approach delegates student-model expertise to large language models (LLMs) while preserving domain expertise for students and teachers, effectively mitigating the risks associated with LLM hallucinations. By abstracting learning into a universal measure, the "Bloom Unit," and enabling scalable, information-driven insights, we provide a foundation for adaptable, inclusive educational tools that transcend domain, age, and cultural boundaries.

The solution's scalability and impact are amplified through mechanisms such as swarm learning, gamification, and the development of a knowledge graph that captures compounded student models. These features foster positive network effects and enable resource-efficient learning pathways, benefiting not only students but also their peers and educators.

By focusing on the first principles of learning, this project seeks to revolutionize, rather than merely evolve, the EdTech landscape. It demonstrates that addressing the kernel of learning can lead to profound, scalable, and transformative outcomes, paving the way for an era of highly personalized, efficient, and effective education for all.

9. Annex

9.1 Research Problem: Characteristics and Assessment

The problem can be stated clearly and concisely: The research investigates the causal or significant correlational relationship between the learning experience—framed within theories such as Bloom's taxonomy, Deep Processing, Transfer Theory, Kahneman's theory, and the Multiple Store Model of Memory—and school variables. These variables include inputs (e.g., LMS, infrastructure, lecturer-to-student ratio) and operations (e.g., lectures, practice sessions, test-taking). Operationalizing these variables draws from personal experience and existing literature.

The problem generates research questions: The study poses numerous primary and subsidiary questions. The challenge lies in asking the right questions at an appropriate level of abstraction—neither too broad nor overly specific. The literature review serves as a guide in refining these questions.

It is grounded in theory: The research is both explanatory (internal validity) and generalizing (external validity), designed within a correlational framework. By adopting a hypothesis-testing approach, it aims to predict plausible correlations based on established theories and prior research.

It relates to one or more academic fields of study: The problem spans multiple disciplines, including educational psychology, developmental psychology, cognitive psychology, behavioral psychology, pedagogy, instructional technology, cognitive science, and artificial intelligence. However, domains like the neuroscience of learning and information theory are excluded for being too abstract for this study's focus.

It has a base in the research literature: The problem synthesizes concerns from various areas into a single statistical model to achieve: 1. Comparative analysis of correlation coefficients across school variables.
2. Factor analysis to identify underlying commonalities influencing the learning experience using techniques like PCA.
To the best of the author's knowledge, this integrated approach has not been applied to single school variable models.

It has potential significance/importance: The problem is driven by the author's personal experience, observational insights, and literature review, making it highly relevant to students' educational needs. Its focus on improving the educational experience ensures its broader significance.

It is do-able within the time frame, budget, or principle: The study is low-cost and feasible within a reasonable time frame, despite being intellectually demanding. Its focus on achievable improvements, even if incremental, ensures practicality.

Sufficient data are available or can be obtained: Initial data scarcity was resolved through a triangulated approach—collecting data from multiple sources and combining datasets. This not only increased the dataset size but also strengthened data analysis and validation.

The researcher's methodological strengths can be applied to the problem: Partially. The author initially faced challenges in academic writing, data analysis (descriptive and inferential statistics), and literature review synthesis. These were addressed through targeted courses (e.g., on Coursera and Scribbr) and self-directed learning in statistics and data science.

The problem is new and not already answered sufficiently: Partially. While the study addresses established questions, it takes a novel, eclectic approach by integrating multiple dimensions of learning and school variables into a single model. The focus is on generating new data and insights rather than addressing entirely unexamined questions.

9.2 Solution Performance

Table 1. Summary of Solution Attributes

Variable	Description
Innovation	- Delegates student-model expertise to LLMs and domain expertise to students/teachers, avoiding pitfalls like LLM hallucinations.
	- Contrasts with competitors: BloomBERT (taxonomy classification), Khanmigo (domain expertise), TahseenAI (behaviorist fallacies).
	- Focuses on Bloom Taxonomy, prioritizing student-centric over domain-specific modeling.
Impact	- Creates a universal measure ("Bloom Unit") across domains, ages, and levels.
	- Enables efficient resource allocation via personalized student profiles.
	- Boosts engagement through gamification, socialization, and avatars.
Feasibility	- Implements Bloom Taxonomy to refine student learning models.
	- Enhances strategies and corrects gaps with actionable insights.
Scalability	- Diffusive, anti-rival, and adaptable across domains and ages.
	- Promotes collaborative and autonomous learning via shared models.
	- Supports swarm learning and knowledge graphs.
Adaptability	- Dynamically adjusts Bloom weights by domain, age, and context.
	- Balances personalization to maintain autonomy and optimize difficulty.

References

[1] Smith, 1979; Lang, 1996. *A Study of Cognitive Development in Education.*

[2] Fleming, 1999. *Learning Styles and Their Impact on Education.*

[3] Lang, 1996; Bligh, 1972. *Teaching Methods and Learning Outcomes.*

[4] Koc, 2015; Ajani and Akinyele, 2014. *A Framework for Educational Development.*

[5] Bayo, 2005. *A Review of Learning Strategies and Their Applications.*

[6] Bloom, 1982. *The Nature of Educational Goals and Their Classification.*

[7] Bloom, 1968. *Taxonomy of Educational Objectives: The Classification of Educational Goals.*

[8] Plomin and Daniels, 'Gloomy Prospect', 1987. *The Role of Genetics in Cognitive Development.*

[9] Soria, 2013. *The Effectiveness of Personalized Learning Strategies.*

[10] Flynn, 1981. *The Flynn Effect: The Impact of the Environment on Intelligence.*

[11] Karpick, 2011. *The Importance of Retrieval Practice in Learning.*

[12] Koc, 2015. *Innovative Educational Methods and Student Success.*

[13] Bolden, 2017. *Reflective Practices in Educational Settings.*

[14] Coe, 2015. *The Evidence for Effective Teaching Practices.*

[15] Pickering, 2017. *The Cognitive Impact of Interactive Learning.*

[16] Herman, 2005. *Cognitive Approaches to Instruction and Learning.*

[17] Cannell, 1988; Linn, Graue, and Sanders, 1989; Shepard, 1989. *Educational Testing and Its Implications.*

[18] Kuncel, 2007. *The Predictive Power of Standardized Testing.*

[19] Rothman, 1988; Popham, 2001. *Assessing Educational Measurement and Evaluation.*

[20] Popham, 2001. *The Case for Performance Assessment.*

[21] Haney and Madaus, 1986. *The Use and Misuse of Testing in Education.*

[22] Campbell, 1976. *The Role of Education in Society.*

[23] Bracey, 1989; Dorr-Bremme and Herman, 1986. *Educational Policy and Assessment.*

[24] Wrightstone, 1967. *Classroom Management and Student Learning.*

[25] Cohen, 1987. *Education and Cognitive Development.*

[26] Popham, 1985. *Educational Measurement: The Issues at Stake.*

[27] Flavell, 1979. *Cognitive Development and Educational Practices.*

[28] Kahneman, 2011. *Thinking, Fast and Slow.*

[29] G. Miller, 1959; H. Simon and W. Chase, 1973. *Human Memory and Cognitive Psychology.*

[30] T. Hyde and J. Jenkins, 1969. *Cognitive Processes in Learning.*

[31] Atkinson and Shiffrin, 1968. *The Human Memory: A Proposed Model.*

[32] Ebbinghaus, 1885. *Memory: A Contribution to Experimental Psychology.*

[33] John Sweller, 1980s. *Cognitive Load Theory and Its Application to Education.*

[34] Perkins and Salomon, 1992. *Transfer of Learning: An Integrated Approach.*

[35] Thorndike, 1923; Scribner and Cole, 1981; Pea and Kurland, 1984. *Theories of Intelligence and Cognitive Development.*

[36] Clements and Gullo, 1984. *Young Children's Cognitive Development and Early Education.*

[37] Lindquist, 1951. *Educational Testing and Measurement.*

[38] Tetlock, 2005. *Expert Political Judgment.*

[39] Fleming and Braume, 2006. *Learning Styles and Educational Success.*

[40] Willingham, 2015; Rogowsky, 2015. *Cognitive Science and Education.*

[41] Willingham and Dobolyi, 2017. *Cognitive Psychology in Education.*

[42] Coe, 2017. *Effective Instructional Strategies.*

[43] Skinner, 1915; Kahneman and Tversky, 1979. *Behavioral Psychology and Decision Making.*

[44] P. Freire, 1964. *Pedagogy of the Oppressed.*

[45] G. Siemens, 2004. *Connectivism: A Learning Theory for the Digital Age.*

[46] Nolen, 1988. *Developmental Psychology in Educational Contexts.*

[47] Craik and Lockhart, 1972; Craik and Tulving, 1975. *Levels of Processing: A Framework for Memory and Learning.*

[48] Boyd, 1986. *Social Learning and Educational Outcomes.*

[49] Bloom, 1956. *Taxonomy of Educational Objectives.*

[50] Anderson and Karathwohl, 2001; Chuches, 2008. *Revised Bloom's Taxonomy.*

[51] K. Collis and J. Biggs, 1982. *Teaching for Quality Learning.*

[52] McGuire, 2015. *The Art of Learning: Mastering Cognitive Skills.*

[53] L. Vygotsky, 1930-1934. *Mind in Society: The Development of Higher Psychological Processes.*

[54] Bloom, 1984. *The Cognitive Domain: Foundations for Teaching.*

[55] Newell, 1990; Anderson, 1993. *Cognitive Psychology: A Student's Handbook.*

[56] Tom, 2024; Schaul, 2024. *Boundless Socratic Learning with Language Games.*

[57] Junior Cedric Tonga and Benjamin Clement and Pierre-Yves Oudeyer ,2024. *Automatic Generation of Question Hints for Mathematics Problems using Large Language Models in Educational Technology.*