Artificial Intelligence in Education: Adapting traditional practices for effective student learning

Thomas Zijl
October 11th, 1999
IJmuiden, The Netherlands

August 13, 2025, 97 pages

Academic supervisor: Dr. A.M. (Ana) Oprescu, a.m.oprescu@uva.nl
Academic supervisor: Dr. A.S.Z. (Adam) Belloum, a.s.z.belloum@uva.nl
Host organisation: University of Amsterdam - Programme Committee

Universiteit van Amsterdam
Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Master Software Engineering
https://www.uva.nl/

Abstract

The rapid emergence of Artificial Intelligence (AI) marks a technological shift with profound implications across multiple sectors, including education. This thesis examines how educational institutions can adapt curriculum design, assessment methods, and teaching practices to remain effective in learning environments where students have access to widely used AI-based tools such as popular Large Language Model (LLM)s like ChatGPT, Gemini, and Claude. Using qualitative analysis of literature, institutional policies, and emerging practices, it explores how AI is reshaping teaching, learning, and assessment. The findings highlight a dual challenge: while AI-driven tools can enhance personalization, automate routine tasks, and foster student engagement, they simultaneously disrupt core educational processes by raising concerns over academic integrity, reducing active cognitive engagement, and undermining the reliability of traditional assessment formats. This study proposes a two-tiered response. In the short term, educational institutions should adapt existing course structures by aligning assessments with demonstrable learning outcomes, integrating AI literacy into curricula, and designing authentic evaluation methods resistant to AI usage. In the long term, a more profound transformation is required: rethinking curriculum design, fostering intrinsic motivation, and developing resilient pedagogical models that balance AI use with the cultivation of enduring human skills such as critical thinking and creativity. The research highlights the urgent need for coordinated policy development, empirical testing, and international collaboration to ensure that the adoption of generative AI supports rather than undermines the fundamental goals of education. It concludes that thoughtful integration of AI presents not only a disruption but also an opportunity to create a more engaging and future-ready educational system.

Contents

Chapter 1

Preliminaries

Extended Glossary

Abbreviation	Term	Definition
UvA	University of Amsterdam	A university in Amsterdam, the Netherlands. The UvA is among Europe's finest universities. With over 40,000 students, 6,000 staff, and 3,000 PhD students, they are an intellectual hub in the world [1].
PC	Programme Committee	The Programme Committee (Dutch: Opleidingscommissie, OC) is a participatory body within the university composed of both students and lecturers. Operating at the local level—typically within a specific programme or group of related programmes—the committee serves as a forum for discussing the quality and organization of education. Within this framework, the OC provides advisory input on curriculum design, quality assurance, and educational policy development [1].
OC	Opleidingscommissie	The Opleidingscommissie or Programme Committee in English. These terms share the same definition.
AI	Artificial Intelligence	AI is a broad field of computer science focused on creating machines or systems that can perform tasks that typically require human intelligence. This is an umbrella term that includes many subfields, like machine learning, computer vision, robotics, and natural language processing.
GenAI	Generative Artificial Intelligence	GenAI is a type of AI that can create new content, such as text, images, music, or code, based on training data. Instead of just analyzing or classifying data, GenAI can generate new, original-like output.
LLM	Large Language Model	LLMs are a specific kind of GenAI model trained on massive amounts of text data to understand and generate human-like language. These models can answer questions, write essays, summarize content, translate languages, and more.

PISA	Programme for International Student Assessment	An international comparative study that tests the skills and knowledge in science, reading, and mathematics of 15-year-olds.
OECD	The Organisation for Economic Co-operation and Development	An international organisation that works to build better policies for better lives. They draw on more than 60 years of experience and insights to shape policies that foster prosperity and opportunity, underpinned by equality and well-being.
AIED	Artificial Intelligence in Education	The field of research dedicated to invesitgate the use of AI in education.
EER	Educational effectiveness research	Research on the effectiveness of education.

List of Figures and Tables

Chapter 2

Introduction

In a constantly evolving world, technological advancements are inevitable, and their impact is felt across every sector. Universities cannot remain passive in the face of these changes. The current state of education is under increasing pressure as accessible LLM-tools can now perform many tasks that were once essential for students to learn and practice themselves. Traditional education models are being challenged, and educators raise questions about their continued effectiveness and relevance. Education must respond adequately to this shifting landscape and redefine its role in developing critical skills and knowledge to ensure educational practices are meaningful. This research contributes to that response by synthesizing literature and seeking answers to significant questions surrounding the integration of AI into education. While it originated within the PC of the Software Engineering program at the University of Amsterdam (UvA), the study takes a broader, global perspective on the future of education in an AI-driven world.

2.1 Artificial Intelligence and education

In late 2022, OpenAI released the first version of ChatGPT, introducing the world to a groundbreaking new technology. The rise of effective LLM technologies shocked the world, as it could answer all your questions and automate tasks once thought only people could do [2]. This technological breakthrough has sparked an unprecedented wave of discussion and research in both academia, business, and policy circles. Within just two months, ChatGPT reached over 100 million users globally [3], showing its rapid adoption and the urgency for educators to respond. Ever since this shift, there has been considerable interest in how LLM technology will impact major sectors, such as education [4][3][2].

Importance of education

We refer to education as a structured and organized process, conducted through institutions such as schools and universities, where individuals learn not only academic content but also social norms, cultural values, and life skills [5][6]. Education is one of the cornerstones of society and a significant systematic way to transfer knowledge between people [6]. Education impacts people both on individual and societal levels [7]. Additionally, education has a direct impact on economic development, quality of life, and social integration, as well as globalization [8]. The journal Globalization, Societies and Education states "Formal education is the most commonly found institution and most commonly shared experience of all in the contemporary world" [9]. Despite differences in style and approach, education has an overall similar structure in many parts of the world, including teacher-centered lectures, in-class assignments, homework, and tests [10].

Technological shift

The impact of advancing technology on education, including LLM technology, is nothing new. However, the breakthrough in LLM technology occurred only recently. As a result, extensive literature on the subject is still emerging, with the first academic papers appearing in early 2023, shortly after the initial release of ChatGPT [3]. Educators, policymakers, and researchers continue to debate how educational institutions can best adapt to this technological shift [4][2]. Students can benefit from a personal virtual assistant that offers instant feedback, on-demand answers, and explanations of complex topics. In addition, these technologies can help enhance language, generate ideas, compose essays, summarize content, translate text, and more. On the other hand, studies show that LLM based tools can also negatively impact students' innovative capacities, learning capabilities, critical thinking, and collaborative learning [11]. These systems can render assignments ineffective by completing tasks entirely on behalf of the student [3]. Additionally, they may provide incorrect information without the user's awareness.

Artificial Intelligence policies at the University of Amsterdam

Currently, the use of commercial LLM technology is prohibited at the UvA. As stated on their public website, "ChatGPT is a product of a commercial company. Currently, it is unclear what is done with user data and the information entered. This makes the tool not (yet) suitable for use within the UvA..." [12]. This is not the only reason for the restriction. Another concern is academic integrity: "As a UvA student, you should expect high-quality and innovative education. This includes an intrinsically motivated attitude. It is still important to do your own writing assignments, and not have a Generative Artificial Intelligence (GenAI) tool do it for you" [13]. UvA regulations state that instructors must be able to assess students' abilities based on their work. Therefore, students are required to submit assignments they have completed themselves, not work generated by others or by AI systems [13]. If students violate this rule, their work may be classified as plagiarism. In severe cases, this can lead to expulsion and the inability to complete their degree at the university: "The UvA takes strict action against plagiarism. Those who commit serious fraud can be deregistered and prevented from completing their studies" [14]. Despite these clear guidelines, it remains uncertain to what extent students follow them. Detecting AI-generated work is complex, and proving its use can be even more challenging. This situation also brings up philosophical questions, for instance: *If an LLM tool improves your work, is it still your own?""* The PC of the master Software Engineering initiated this thesis topic, which has evolved into a study addressing these issues on a global scale rather than being limited to the UvA context.

Enforceability of Artificial Intelligence related policies

The UvA is actively working to improve its AI policies and explore the potential of AI in education. As stated on their website: "Developments in the field of AI move extremely fast, which is why the UvA regularly reviews its AI policy. We are interested in the potential of AI programs in the fields of education and research. In the development of our AI policy, the current focus is on the role of AI in education. [12]". The UvA is taking concrete steps by establishing a dedicated task force, developing its own LLM tools, and collaborating with PhD students to create more effective AI detection systems [15][12]. They also acknowledge the challenges faced by educators: "For a large group of lecturers, the sudden emergence of ChatGPT has created extra work, mainly because they have had to adjust their assessments to prevent fraud. Accordingly, in the short term, the UvA policy should focus on supporting degree programmes in general and lecturers in particular in dealing with AI in their teaching and assessment" [12]. Most importantly, the UvA is examining how to integrate GenAI into traditional educational settings. This highlights the importance of ongoing research in this area: "For all education offered by the UvA, it will be necessary to assess whether programmes' exit qualifications are still appropriate for this new reality. Do the current exit qualifications still allow us, educators, to prepare people for the current and future labour market? Or do they need to be refined?" [12].

2.2 Problem statement

One of the main concerns is the difficulty in detecting LLM use directly from the source material. Even if there is suspicion that a student has used GenAI, it is often challenging to prove [16]. Other factors—such as writing style, prior knowledge, or external assistance—could also explain suspicious work, and existing LLM detectors are prone to false positives [4][16]. At present, implementing a complete or partial ban appears to be highly challenging. For example, in the case of the UvA's policy [12], the ban is essentially unenforceable, which renders it largely ineffective. This lack of enforceability creates the risk that students may disregard the rules without consequence, resulting in an unfair distribution of workload and potentially leading to the graduation of students with fewer essential skills. Given the considerable potential of LLMs in education [17][3], and the impracticality of strict bans, the focus should perhaps shift toward policy reform and student learning. Rather than resisting the presence of GenAI, institutions may need to rethink traditional educational approaches to accommodate this technology.

2.3 Research questions

To address the challenges posed by the integration of LLM in education, this research explores how traditional educational approaches can be adapted to incorporate LLM technology, while ensuring that students continue to perform and learn as effectively—or even more effectively—than before.

This study aims to answer the following main research question:

> **RQ:** How can educators in higher education adapt traditional course practices to maintain intended student learning outcomes while permitting the use of Large Language Model technology by students?

This main research question is supported by the following sub-questions:

- **SRQ1:** What characteristics make traditional course settings effective as an educational method?
- **SRQ2:** How are traditional course components affected by the emergence of Large Language Models?
- **SRQ3:** What strategies can educators adopt to ensure meaningful student learning in the presence of Large Language Model technology?

The main research question encompasses four key components: the actors, referring to educators particularly involved in course delivery at Dutch universities; the action, which involves redesigning traditional courses; the goal, defined as achieving the desired student learning outcomes; and the context, which we define by the ongoing technological shift brought about by the emergence and increasing use of LLMs in education.

To clarify the research question further:

- *"Maintain intended student learning outcomes"* refers to identifying which changes are necessary to ensure students still acquire essential skills, even when using LLMs.
- *"Educators"* refers to lecturers and teaching staff directly involved in course delivery within higher education institutions, especially in the Dutch context.
- *"Learning outcomes"* are defined as the measurable differences in a student's knowledge or skills before and after completing a course, typically assessed through exams or other evaluations.
- *"Traditional courses"* are characterized by one or more lecturers delivering monologic lectures, supplemented with assignments, homework, and standard testing.
- *"Allowing"* implies acknowledging the widespread use of LLMs by students and adapting educational practices accordingly, recognizing that detection is difficult and proof of misuse even harder. For example, this might involve allowing LLMs to be used while ensuring that students are still acquiring the necessary skills.
- *"LLM technology"* includes LLM tools such as ChatGPT, Claude, Gemini, DeepSeek, and others.

2.4 Context of the problem

This paper focuses on the integration of LLMs in educational settings for students aged 18 to 25, a demographic that primarily includes those in post-secondary or higher education. The thesis examines how educational systems can adapt to the increasing use of LLMs by students without external limitations. We pay particular attention to curriculum design, assessment strategies, and institutional policies. To provide a comprehensive understanding of the topic, the literature review will draw from a broad range of global academic sources.

2.5 Addressing the problem

This research examines the integration of AI tools, particularly those based on LLMs, into higher education, to ensure that students continue to develop essential skills and maintain academic integrity. While AI holds impactful potential, its application in educational settings raises concerns regarding academic standards, instructional effectiveness, and the risk of misuse [16]. This study aims to provide practical guidance for educators on how to incorporate AI in ways that foster critical thinking while mitigating risks such as academic dishonesty and over-reliance on technology. It fills a gap in the existing literature by offering actionable recommendations that strike a balance between technological innovation and pedagogical objectives. Additionally, it provides an overview of current frameworks, guidelines, and institutional approaches. While much of the existing research emphasizes ethical and theoretical perspectives on AI, this study focuses on its practical application in real-world teaching contexts to enhance student learning and engagement.

2.6 Contributions

This research makes the following contributions:

1. It presents a systematic literature review on the integration of LLMs in education.
2. It offers practical, evidence-based guidance for educators on how to redesign traditional courses to accommodate LLM use, grounded in existing research and empirical insights.
3. It provides a comprehensive overview of current developments in the Artificial Intelligence in Education (AIED) landscape, including common guidelines, institutional policies, and the impact of AI on traditional educational practices.

2.7 Outline

Chapter 3 provides the background and context of the study. Chapter 4 reviews existing literature and related work on the topic. Chapter 5 outlines the research methodology. Chapter 6 presents the findings, drawing on literature and data to address the main research question. Chapter 7 interprets and discusses the results in relation to the research objectives. Chapter 8 summarizes the main conclusions and reflects on the broader implications of the study. Additionally, Chapter 10 provides an extended, annotated bibliography with evaluations and summaries of key sources to support easier referencing and future reuse.

Chapter 3

Background

This paper examines how educators can adapt traditional educational practices in response to the technological advancements introduced by LLMs. Before addressing this core question, it is essential to establish the background of the research first.

3.1 The historical perspective

The term "Artificial Intelligence" was first coined at the 1956 Dartmouth workshop, defined as "the science and engineering of making intelligent machines, especially intelligent computer programs" [18]. Technological advancements challenge established practices, necessitating a critical evaluation of both the potential advantages and limitations associated with emerging technologies [19]. Similar questions about GenAI were raised during the widespread adoption of Google, as researchers pondered its impact on education [20]. Tools like ChatGPT hold promise for enhancing learning by delivering personalized feedback, tailored explanations, and immersive virtual simulations that support experiential learning. However, it is equally important to critically examine the limitations and potential challenges associated with these technologies [17].

While the origins of AI date back several decades, initially centered on early systems and explorations of the relationship between AI and human cognition, this paper focuses on more recent developments. Particularly with emphasis on the significant progress made after 2020.

3.2 The technological perspective

Over the past decade, significant advancements have been achieved in the field of AI, mainly driven by machine learning techniques. Broadly defined, AI refers to systems that can perform tasks typically requiring human intelligence, such as mimicking reasoning, perception, and decision-making. Recent trends in AI emphasize GenAI, a subset of AI that focuses on creating new content, such as text, images, or code, as well as continued progress in reinforcement learning, multi-modal AI (which integrates text, images, and sound), and ethical AI governance.

One of the most prominent developments within genAI is the rise of LLMs, such as ChatGPT, which are capable of generating coherent, contextually appropriate text based on vast amounts of training data. These models have brought AI from its theoretical roots to practical systems that exhibit creative and intelligent behavior across various domains [18].

These technological developments have laid the foundation for the intersection of AI with education. The rise of AI has been accompanied by the emergence of AIED as a distinct research field. AIED is an interdisciplinary domain that investigates how intelligent technologies can support and enhance teaching and learning processes [3].

3.3 The education perspective

Education consists of two major components: what is taught (curriculum) and how it is taught (pedagogy). To understand how AI impacts education, it is essential first to clarify both of these concepts. Curriculum design involves the deliberate planning and sequencing of learning experiences to achieve defined objectives. A curriculum typically specifies what students should learn and how it is organized [21]. Pedagogy, on the other hand, refers to the methods and strategies of instruction, essentially describing how educators teach and how learners engage. It focuses on the techniques teachers use and the learning activities students undertake [5]. Foundational pedagogical theories range from behaviorist models, such as drill-and-practice, to constructivist approaches, where learners build knowledge through experience. However, the core aim of most educational institute is holistic skill development [8]. Modern curricula extend beyond basic literacy and numeracy to encompass a range of competencies, including critical thinking, problem-solving, collaboration, creativity, and digital literacy, that learners need for success in the 21st century [18]. Effective education requires the alignment of curriculum goals and pedagogical practices that point toward the holistic goal of personal development [22]. Any integration of AI into education must respect these foundations of curriculum, pedagogy, and the overarching mission to develop well-rounded skills in learners.

There are two primary pedagogical approaches in education: inductive and deductive teaching. Traditional engineering and scientific education often employs a deductive approach, which begins with general principles, derives models from these principles, and then applies them in structured exercises, such as homework assignments and exams. Conversely, inductive teaching starts with specific observations, case studies, or real-world problems. Students analyze these situations, creating a natural need for relevant principles, which are subsequently introduced or discovered through guided instruction. This approach promotes active learning and the development of critical problem-solving skills [23].Inductive and deductive teaching represent two overarching educational paradigms, each encompassing a wide range of corresponding practices for both teachers and students [24]. This paper examines how AI reshaped these approaches and their associated practices.

Note on Artificial Intelligence in Education

Numerous organizations are actively researching the current state of AI [10, 25], with Stanford's AI-index emerging as a leading contributor in this field. Understanding the capabilities of current AI tools is essential for informing how existing practices, particularly in education, can be adapted. Recent benchmarks show that AI systems have surpassed human performance in areas such as image classification, visual reasoning, and English comprehension. However, they still underperform in more complex domains, including competition-level mathematics, visual commonsense reasoning, and strategic planning [26][27].

The assumption that AI systems are widely used in education, particularly among students, is supported by several studies. Overall, the studies indicate that students hold a favorable view of GenAI technologies and express a desire to incorporate tools like ChatGPT into both their learning activities and future professional endeavors. Most students report using LLM based tools in a medium to high capacity for educational purposes [16]. Chan and Hu highlights significant concerns among students regarding AI usage: “"Unlike willingness, descriptive statistics show that students expressed a slight favor of concerns about GenAI. They expressed the least positive opinions about whether people will become over-reliant on GenAI technologies ... the highest rating was for how these technologies could affect the value of university education..."” [16]. These findings underscore the importance of addressing how AI can be effectively incorporated into educational settings while preserving an environment that fosters meaningful and effective learning for students.

Chapter 4

Related work

The landscape of AI in education shifted with the public release of ChatGPT in late 2022. This development became an incentive for progress in research on AIED from early 2023 onward [3]. As a result, a rapidly growing body of literature now addresses the educational implications of GenAI, even though the topic has emerged in its current form only within the past two years. This thesis specifically explores how educators can adapt traditional education systems in response to students' increasing use of LLMs. While some recent studies explore similar themes, they often approach the topic from different perspectives. For example, re-evaluations of existing educational paradigms in light of GenAI [11], assessments of the current global state of education [10], domain-specific analyses such as the advantages and disadvantages of ChatGPT in engineering education [28], and broader discussions on the impact of tools like ChatGPT on contemporary teaching methods [3].

This chapter will delve further into the relevant literature, highlighting current discourse and identifying key contributions. It will also clarify the specific gap that this research aims to address.

4.1 Related papers on Artificial Intelligence in education

LLM technology provides numerous opportunities for current education practices. The current and potential roles of these tools are presented by Labadze *et al.* in "Role of AI chatbots in education: systematic literature review." His findings highlight several benefits of integrating AI-powered chatbots into educational settings, from both student and educator perspectives. According to the study, students primarily benefit in three key areas: assistance with homework and studying, personalized learning experiences, and the development of diverse skills. For educators, the primary advantages lie in time savings and pedagogical improvement. However, major challenges remain, particularly regarding reliability, accuracy, and ethical considerations [2].

With the emergence of increasingly advanced LLM technologies, educators must reconsider how they design and assess course curricula. If a LLM can instantly produce a passable essay or solve problem sets, traditional forms of assessment may lose their discriminative power. Early research has warned that conventional take-home essays or programming tasks, once considered reliable indicators of student understanding, may no longer provide valid measures of individual ability in contexts where AI assistance is readily available [16].

Research has shown that tools like ChatGPT exhibit human-like capabilities, though these vary across subject domains. Although LLMs hold considerable potential to function as teaching assistants or virtual tutors, their reliability remains inconsistent. More importantly, there is still a lack of long-term, in-depth research on the use of these tools in education, and the technology itself continues to evolve. For this reason, proactive measures are necessary to redesign assessment methods and institutional policies, ensuring that educational systems are prepared to respond to ongoing developments in the years to come [3].

Similarly, Albadarin *et al.* investigates how both students and educators have used a widely available LLM tool, ChatGPT, and examines the outcomes of this use. In "A systematic literature review of empirical research on ChatGPT in education," the authors explore diverse applications of ChatGPT in educational settings, offering a more detailed understanding of both the advantages and limitations of LLM integration in learning and teaching practices [11].

There is a need for engineering educators to grasp the implications of GenAI and to explore ways to adapt education system so that future workers can benefit from the technology while mitigating potential risks [28]. Although this specific paper focuses on the prospective role of ChatGPT in education, it provides limited insight into the concrete methods of implementation. It does not thoroughly address the connection between GenAI use and educational policy.

The use of LLMs in education has caused mixed reactions among students and educators [29]. In response, researchers began to explore the potential applications of LLMs, aiming to provide clarity and guidance for educators on their use. However, similar to other studies, this research also lacks a detailed discussion on how educational policies and curricula could or should evolve to effectively integrate LLM technology in a way that enhances learning outcomes.

The larger part of the literature on LLM use in education has been exploratory and descriptive. While researchers did conduct some case studies, much of the existing work consists of conceptual papers or commentaries that hypothesize potential benefits. Beltozar-Clemente *et al.*, for instance, evaluated LLM performance on academic tasks to assess its potential as a tutoring tool. Their study tested ChatGPT on university entrance exam questions to determine whether it could function as a competent tutor. The results indicated that while ChatGPT was able to answer many multiple-choice questions correctly, it struggled with more complex, higher-order problems, highlighting the current limitations of the technology [30].

4.2 University policies

While advancements in AI offer considerable opportunities to enhance educational experiences, they also raise serious concerns, particularly regarding academic integrity [4]. This reflects the broader context in which AI is being introduced to education: a powerful and promising technology accompanied by complex challenges. New literature provides a detailed examination of educators' perspectives on LLM use by analyzing the academic policies and guidelines implemented across various U.S. universities [4]. These policies specifically address how AI tools, such as ChatGPT, are being incorporated into teaching and learning practices [4]. The findings indicate that most institutions are adopting a cautious stance toward GenAI. Key areas of concern include ethical use, accuracy, and the protection of personal data. In response, many universities have taken proactive measures by offering a range of support resources, such as syllabus templates, workshops, curated articles, and individual consultations, covering topics from technical fundamentals and ethical considerations to pedagogical applications, data privacy, and detection tools [4]. This approach aligns with the UvA, which has adopted a similarly "open but cautious" stance regarding the use of GenAI in education [12]. The same goes for other (international) universities [31][32][33][34].

4.3 The research gap

Direct and well-informed action is needed to adapt education in response to the rise of AI tools. This action includes updating assessment methods and institutional policies to ensure they remain relevant and enforceable [3]. Beyond methodological changes, there must also be a shift in mindset and an increase in knowledge among both students and educators. This shift will enable them to use these tools responsibly and follow university regulations. Many students and instructors express anxiety and hesitation regarding the integration of AI in educational settings, a response that deserves to be acknowledged and addressed [4]. This response should come in the form of clear, practical guidelines, targeted instructor support, and redesigned curricula that reflect the realities of AI-enhanced learning environments [11]. Although these needs are increasingly recognized, research in this area is still in development, and there is broad agreement on the importance of proactive collaboration between policymakers, educational stakeholders, and practitioners to improve the situation cite albadarin2024. This thesis contributes to that ongoing conversation by offering concrete insights and recommendations for how traditional education systems can meaningfully and responsibly integrate LLM technologies.

Chapter 5

Methodology

This chapter outlines the qualitative methodology used to investigate how higher education can adapt traditional teaching practices to integrate LLM technology while maintaining effective student learning outcomes. The primary challenge of this research lies in integrating diverse fields; educational practices and AI.

5.1 Research methods

Defining the research methods is a crucial step before collecting and analyzing data. This section focuses on two key aspects of the design: the method of data collection and the method of data analysis. This study will explore existing research and synthesize it to generate new insights, in a primarily qualitative approach. The primary method will be a literature review. The research relies on secondary data, which facilitates the synthesis of existing knowledge and information. While secondary data is accessible and spans extensive datasets and literature, it requires careful processing to ensure suitability for analysis, as there is no control over how the data was initially generated.

How can educators in higher education adapt traditional course practices to maintain intended student learning outcomes while permitting the use of Large Language Model technology by students?

- Approach: By synthesizing insights from peer-reviewed literature, educational frameworks, empirical case studies, and grey literature, we can formulate an overview of evidence-based design principles that support the integration of LLMs into traditional education without compromising student learning outcomes.
- Hypothesis: Educators can adapt traditional courses by redefining assessment practices, emphasizing higher-order cognitive skills, and rethinking instructional design to include productive LLM use.
- Data sources: We address the main question through the answers to the following sub-questions.

What characteristics make traditional course settings effective as an educational method?

- Approach: By reviewing foundational and pedagogical literature on effective traditional teaching methods, we can identify core characteristics that contribute to student learning in conventional classrooms.
- Hypothesis: Traditional courses are effective when they provide structured learning, guided practice, formative feedback, and well-aligned assessments that target specific learning objectives.
- Data Sources: Peer-reviewed literature on educational psychology and pedagogy, Meta-analyses of instructional effectiveness, Reports from national education bodies.

How are traditional course components affected by the emergence of Large Language Models?

- Approach: By analyzing case studies, foundational pedagogical literature, policy reports, and institutional documentation that describe LLM use in education, we can identify which traditional course components are most impacted by the availability of GenAI tools.
- Hypothesis: LLMs most significantly affect assignments and assessments that provide time and room for generative outputs, such as essays and problem sets, potentially undermining authenticity unless instructional design is adapted.
- Data Sources: Grey literature (e.g., government documents, institutional reports), Peer-reviewed studies on LLM integration, Surveys or datasets showing student usage patterns

What strategies can educators adopt to ensure meaningful student learning in the presence of Large Language Model technology?

- Approach: If we examine literature on academic integrity, AI literacy, formative assessment design, and student engagement, we can derive actionable strategies for teaching with LLMs that maintain or enhance learning outcomes.
- Hypothesis: Educators can maintain meaningful learning by integrating AI literacy into curricula, redesigning assessments to focus on critical thinking, and using LLMs as scaffolding tools rather than answer engines.
- Data Sources: Educational technology and instructional design literature, Case studies and pilot programs involving LLM use, Grey literature on policy and AI literacy education.

5.2 Validation

We validate this study through expert review, comparative analysis with existing research, and a potential real-world implementation. Experts will provide feedback to evaluate the feasibility and relevance of the proposed approach. The findings will also be compared with existing studies to ensure consistency. Possibly, A field experiment may be conducted to assess the practical effects of integrating LLM technology into educational settings. These combined validation methods will ensure the study's credibility and practical applicability. c

Chapter 6

Results

This section presents the results necessary to address the main research question: How can educators in higher education adapt traditional course practices to maintain intended student learning outcomes while permitting the use of Large Language Model technology by students?. To provide a structured and comprehensive response, we divided the question into key thematic areas, each of which is explored in a separate subchapter. Each subchapter examines relevant theoretical frameworks, definitions, foundational concepts, and additional grey literature related to developments in the field of AIED. The analysis considers both the current era of AI and the educational principles that preceded it. For example, in order to assess the impact of LLMs on education, it is first essential to understand the core principles traditionally associated with effective teaching and learning, prior to the widespread availability of advanced AI tools.

6.1 Traditional course settings

The term "Traditional course settings" in educational literature refers to conventional, classroom-based learning environments characterized by face-to-face, teacher-led instruction in a physical location. In these settings, students and instructors are co-present at scheduled times, typically within schools or university classrooms. Although digital tools may supplement such environments or can be joined remotely by students online, the core structure remains grounded in in-person interaction. Defining this concept is important, as this thesis investigates how LLM technology may impact these traditional settings within global higher education.

Courses in this model are typically pre-designed and follow a fixed curriculum, which may include homework assignments, group projects, individual assessments, and examinations (See Table 6.1). Instruction is primarily teacher-centered, with educators directing the pace and flow of information. Key features include a structured timetable, in-person supervision, and a uniform learning pace applied to all students. As such, students' time, place, and pace of learning are mostly predetermined and consistent throughout the course [35].

This approach to education was the prevailing standard globally from approximately 1970 to 2020. Despite numerous innovations after 2000, traditional educational settings from 1970 to 2020 adhered mainly to conventional models of lecture-based instruction, semester-based course structures, predetermined curricula, and exam-focused assessment. The term "traditional" in this context specifically refers to the dominant mode of instruction prior to the widespread availability of GenAI tools, which began to significantly influence education practices after 2022 [3].

6.2 Components of courses

Courses in higher education can encompass various types and quantities of components, including assignments, homework, and examinations. The composition of these components can vary by course, institution, and country. Based on a review of current literature, we present an overview of the most common types of course components. Throughout this paper, the term course components will refer specifically to these identified types.

Overview of course components [27][36][35]

Category	Type	AI Capability (Student Use)	AI Detectability (Teacher Monitoring)	Notes
Instructional Components	Lectures	N/A	N/A	Usually provided by a human instructor
	Seminars	N/A	N/A	Usually provided by a human instructor
	Tutorials	N/A	N/A	AI can partially generate tutorials
	Lab Work	Partially	N/A	Often requires physical, hands-on participation
	Fieldwork	No	No	Requires physical presence and observation
Assessment Components	Examinations (written)	No	Partially	AI can generate written responses; proctoring tools are available
	Examinations (Oral)	No	Partially	AI can facilitate cheating or AI can assist with cheating
	Presentations	Partially	Partially	AI can assist in creation
	Assignments	Yes	Partially	Text-based tasks can be done by AI; plagiarism checks semi apply
	Projects	Partially	Partially	Depends on scope and type of project
	Portfolios	Yes	Partially	AI can help compile content; originality can be partially checked
	Peer Assessment	No	N/A	Requires student interaction
Learning Activities	Group Work	Partially	Partially	AI can mimic input; collaboration is difficult to track
	Simulations	Partially	Partially	AI can power simulations and adapt learning paths
	Research	Yes	Partially	AI can assist in writing and generate complete papers
Homework	Overall	Partially	Partially	Depends on homework type
	Math	Yes	No	Unless highly advanced, most tasks can be completed by LLMs
	Programming	Yes	Partially	Unless advanced, most tasks can be completed by LLMs
	Writing	Yes	Partially	Semi detectable, hard to proof
	Languages	Yes	Partially	LLMs perform well on translation
Learning Outcome Support	Learning Objectives	N/A	N/A	AI can help formulate outcomes
	Rubrics	N/A	N/A	AI can help generate rubrics based on existing ones for similar courses
	Feedback	N/A	N/A	Automated feedback is possible

Table 6.1: Overview of course components and how they relate to AI capabilities

*N/A = Not Applicable

This table outlines standard course components and assesses the extent to which AI tools can perform tasks typically assigned to students, as well as the ability of educators to monitor such use. Drawing on recent research, it highlights where generative AI can assist or automate educational tasks and where detection of AI involvement is possible [37][27]. The table reveals that AI is particularly capable in text-based and repetitive tasks, such as assignments, programming, and writing, although monitoring its use remains limited. In contrast, components requiring physical presence, interaction, or real-time thinking are less affected. This overview supports a clearer understanding of where educational practices may need to adapt in response to AI integration.

6.3 Fundamentals of human learning

To understand how education systems can adapt to new challenges regarding LLM technology, it is essential to consider the foundational principles of human learning. While a comprehensive exploration of the entire learning process is beyond the scope of this thesis, a basic understanding of key learning mechanisms is helpful. These principles underpin effective educational design and practice and are widely recognized in the literature on teaching and learning [24].

Human learning is inherently constructive: learners build upon their prior knowledge, worldviews, experiences, and interests. Learning does not begin from a blank slate; rather, it involves transforming and reorganizing existing mental models, particularly when applying knowledge in new contexts [38]. Teachers play a pivotal role in activating this prior knowledge, helping students integrate new information into their existing mental framework or adjust their existing understanding. Effective learning is situated in context and promotes transfer, the ability to apply knowledge and skills across varied and novel situations. This transfer is most successful when instruction emphasizes core principles, coherent conceptual frameworks, and the deep structure of subject matter [38].

Developing an in-depth understanding and expertise requires engagement with detailed, well-structured knowledge over time. It combines the accumulation of factual content with opportunities to apply knowledge in practical, meaningful ways. Such an application reinforces understanding and supports the formation of new cognitive patterns. To guide this process effectively, teachers themselves must possess both subject-matter expertise and a strong grounding in pedagogy. Their professional development should be continuous and informed by theories of learning and knowledge construction. Creating effective learning environments requires alignment with educational goals, learner diversity, and the evolving demands of society and technology. Students must be intrinsically motivated and understand the relevance and utility of what they are learning. Pedagogical approaches should be both culturally responsive and learner-centered. Assessment plays a critical role—not merely as a tool for evaluation, but as a means to inform instruction and address individual learning needs [38].

Supportive environments are also essential. These include positive social dynamics, collaborative learning opportunities, and connections with families and communities. The integration of technology offers significant potential to enhance learning, but its use must be strategic. Teachers require appropriate training and resources to incorporate digital tools effectively, and the broader cognitive, social, and organizational implications must be carefully considered [22][38].

6.4 Play as a tool for learning

A growing body of evidence challenges the traditional dichotomy between "playful" and "serious" learning. Instead, researchers increasingly emphasize that play is a natural and powerful mode of learning that supports the development of skills relevant to both school and life [39][40]. This perspective is particularly relevant in light of the current shift brought about by advances in AI. In both the short and long term, adapting educational practices to this shift could benefit from integrating more playful approaches. Automating certain repetitive tasks may free up time for alternative forms of learning. Moreover, AI tools can support learning environments where play becomes a central component, making education more engaging, intrinsically motivating, and meaningful.

We can identify five characteristics of high-quality playful learning: it is joyful, meaningful, actively engaging, iterative, and socially interactive [41]. These qualities create conditions in which humans, but specifically children in most research, can develop a deep conceptual understanding and broad "learning-to-learn" skills, such as collaboration, communication, creativity, critical thinking, confidence, and content knowledge. These learning skills are an essential part of effective education [22]. We refer to these conditions as the "6 C's". Far from being a distraction, play nurtures cognitive and socio-emotional growth in tandem, making it one of the most efficient ways to prepare learners for a dynamic and uncertain future [41].

Much research expands on this, arguing that play-based approaches are not limited to early childhood education but also belong firmly within primary school or higher educational settings [39]. Their research synthesizes evidence from a wide range of pedagogies, such as inquiry-based learning, problem-based learning, project-based learning, and Montessori education. It shows that these approaches, when implemented playfully, can significantly enhance learning outcomes [42][39]. We can categorize this into four core dimensions: the learner's experience, the facilitator's role, the design of the activity, and the learning outcomes. The central theme is the importance of student agency; allowing learners to make choices, take risks, and shape their learning pathways contributes to greater engagement and deeper learning [39].

An important framework for putting these findings into practice is the "play-for-learning" model, which enables play to achieve predetermined educational goals [40]. This model highlights the importance of transcending the outdated belief that imagination and reality, or play and learning, are separate and incompatible. Imagination is not a distraction from real learning, but rather an integral part of it [40][39]. Play does not need to be a break from learning—it is learning.

Given these insights, a natural next step is to consider how emerging technologies, particularly AI and LLMs, can support and extend playful learning. AI has the potential to transform traditional learning environments into dynamic, responsive, and engaging spaces. For example, AI-powered platforms can personalize play scenarios to match learners' interests and developmental stages, ensuring that tasks remain both challenging and enjoyable [43]. AI can also act as a scaffold, providing hints, prompts, or feedback in real-time, mirroring the role of an attentive teacher or peer [44]. Additionally, GenAI tools can be used as creative partners in storytelling, art, music, or design, allowing learners to express themselves in novel ways and turning imaginative ideas into tangible outcomes [30][45]. AI based tools can also facilitate playful assessment by unobtrusively tracking learner behavior and offering insights into growth in areas like problem-solving, perseverance, or empathy, skills often neglected in standardized tests but intensely cultivated through play [46].

The integration of AI into education should not aim to replace human-centered practices, such as play, but rather to enhance and extend them. As research shows, learning through play is compatible with deeper cognitive development [39]. When combined with the adaptive, generative, and interactive capacities of AI, playful learning holds even greater promise as a tool for cultivating curiosity, resilience, and creativity.

6.5 Effective education approaches and learning theories

To assess the impact of LLMs on education, it is essential first to understand the foundational principles that have traditionally been associated with effective educational practices. This section focuses specifically on the period between 2000 and 2020, before the emergence of widely accessible LLM technology. While there is ongoing debate regarding the overall effectiveness of current educational designs [47], this study will not delve deeply into that discourse. Instead, we will provide an overview of key pedagogical paradigms, including traditional instructional approaches, constructivist methodologies, collaborative and social learning models, technology-enhanced learning approaches, personalized and student-centered practices, and critical and transformative approaches. In the following chapter, we will also examine the Visible Learning framework, which synthesizes elements from several of these paradigms. This model provides a meta-analytical perspective on what works best in education, grounded in empirical evidence.

Traditional approaches

Traditional teaching methods are foundational to formal education systems worldwide. These approaches are mostly teacher-centered, where the instructor serves as the primary source of knowledge and authority in the classroom. Lecture-based learning is characterized by the verbal transmission of information from teacher to student. Similarly, direct instruction involves a structured and sequenced format, with clear objectives and measurable outcomes. The Socratic method, though older, remains influential; it fosters critical thinking through questioning and dialogue [48][49][10][24].

The table below outlines the most common types of traditional educational approaches found in academic literature.

Type	Description
Lecture-based learning and teacher-centered	Information is delivered verbally to students.
Direct Instruction	Structured, sequenced, and led by the teacher with clear learning objectives.
Socratic Method	Dialogue-based teaching through questioning to stimulate critical thinking.
Drill and Practice	Repetitive practice for skill mastery (common in math and language learning).

Constructivist approaches

Constructivist approaches prioritize the learner's active role in knowledge construction. Inquiry-based learning encourages students to ask questions, investigate solutions, and build understanding through exploration. Problem-Based Learning (PBL) and Project-Based Learning (PjBL) emphasize real-world relevance, requiring learners to solve complex issues or create meaningful projects collaboratively. Discovery learning promotes individual experimentation and hypothesis-testing, while experiential learning incorporates reflection on direct experiences such as fieldwork or simulations [48][49][10][24].

The table below outlines the most common types of Constructivist approaches found in academic literature.

Type	Description
Inquiry-Based Learning	Students explore questions and problems, constructing their understanding.
Problem-Based Learning	Learners solve complex, real-world problems collaboratively.
Project-Based Learning	Students investigate and respond to authentic tasks over extended periods.
Discovery Learning	Students learn through exploration and experimentation.
Experiential Learning	Learning through reflection on doing (e.g., internships, labs, fieldwork).

Collaborative and Social Learning approaches

Collaborative learning emphasizes the social nature of knowledge construction. Cooperative learning involves structured group work where students depend on one another to achieve shared goals. Peer teaching and learning enable students to assume instructional roles, reinforcing their understanding while supporting their classmates. Learning circles provide a more formal structure for group discussions, often centered on literature, ethical dilemmas, or thematic content, and are particularly effective for cultivating inclusive and participatory classroom cultures [48][49][10][24].

The table below outlines the most common types of Collaborative and Social learning approaches found in academic literature.

Type	Description
Cooperative Learning	Students work in small groups with shared goals.
Peer Teaching / Peer Learning	Students instruct or support each other's learning.
Discussion-Based Learning	Emphasis on group dialogue to deepen understanding.
Learning Circles	Structured group discussions focused on specific themes or texts.

Technology-enhanced approaches

Advancements in digital technology have significantly impacted the delivery of education, enabling a range of innovative teaching and learning options. Blended learning combines traditional face-to-face instruction with online components to create a comprehensive learning experience. The flipped classroom model inverts conventional instructional sequences, with students first engaging with new content outside of class—often through video lectures—before applying it during in-person sessions. Gamification specifically introduces game mechanics into learning environments to enhance motivation and engagement, while adaptive learning systems leverage data and AI to personalize content according to each learner's needs and progress [48][49][10][24].

The table below outlines the most common types of technology-enhanced approaches found in academic literature.

Type	Description
Blended Learning	Mix of face-to-face and online learning experiences.
Flipped Classroom	Students learn new content at home (often via video) and apply it in class.
Online / E-Learning	Entirely digital delivery of content and interaction.
Gamification	Using game elements (e.g., points, levels) to increase engagement.
Adaptive Learning	Technology-driven personalization of learning paths and content.

Personalized and student-centered approaches

Student-centered teaching prioritizes the interests, needs, and individual pacing of learners. The Montessori method promotes autonomy through self-directed activity within carefully prepared environments, often with mixed-age classrooms. The Reggio Emilia approach, focused on early childhood education, encourages children to express themselves and explore through art, dialogue, and environmental interaction. Waldorf education emphasizes holistic development, balancing intellectual, artistic, and practical learning in rhythm with the child's developmental stages. Competency-based learning allows learners to progress upon demonstrating mastery, shifting the focus from time-based progression to outcome-based evaluation. Self-directed learning goes a step further by granting students full agency over their learning objectives, resources, and evaluation criteria [48][49][10][24].

The table below outlines the most common types of student-centered approaches found in academic literature.

Type	Description
Montessori Method	Self-directed, hands-on learning with mixed-age classrooms.
Reggio Emilia Approach	Emphasizes exploration, expression, and community in early childhood.
Waldorf Education	Holistic development focusing on imagination, creativity, and rhythm.
Competency-Based Learning	Progress based on demonstrated mastery, not seat time.
Self-Directed Learning	Learners set their own goals and evaluate their progress.

Critical and transformative approaches

Culturally responsive teaching acknowledges and incorporates students' diverse cultural backgrounds into the curriculum, promoting equity and inclusion. Critical pedagogy encourages learners to question dominant narratives and power structures, fostering critical consciousness and civic responsibility. Social-emotional learning integrates emotional and interpersonal skills into academic instruction, recognizing that emotional intelligence is fundamental to student success [48][49][10][24].

The table below outlines the most common types of critical and transformative approaches found in academic literature.

Type	Description
Culturally Responsive Teaching	Validates students' cultural backgrounds in the learning process.
Critical Pedagogy	Encourages learners to question and challenge power structures.
Social-Emotional Learning	Integrates emotional intelligence with academic learning.
Humanistic Education	Emphasizes personal growth, empathy, and self-actualization.

6.6 Effective education and quality learning

To ensure quality learning, we should look at the principle of teaching according to how students learn [50]. Key elements of effective education are increasingly being influenced by AI-related technologies. To respond appropriately, we must first understand the core components of effective education. Only then can we assess how these components are affected by LLM technology and determine how educational practices should adapt in response.

Meaningful learning occurs when students are actively engaged in constructing knowledge, making connections, and applying ideas in context [22]. This deep learning approach contrasts with surface learning, where students rely on rote memorization and aim only to meet minimum assessment requirements. In popular literature, this is illustrated through two student profiles: Susan, a deep learner, and Robert, a surface learner. Susan engages thoughtfully with course content, makes connections across topics, and reflects on her learning process. Robert, by contrast, focuses on memorizing isolated facts to pass exams, often without understanding their relevance. These personas help educators visualize the practical effects of their instructional strategies. When courses are poorly aligned or overly focused on summative testing, they tend to encourage Robert's surface-level behaviors. Conversely, well-designed, aligned learning experiences—like those Susan engages with—foster intrinsic motivation and deeper learning outcomes [50].

This understanding leads to one of the core ideas in effective education: constructive alignment. Constructive alignment refers to designing courses so that Intended Learning Outcomes, teaching activities, and assessment tasks are all in sync [50]. This model is especially relevant in AI-rich educational settings, where students may be tempted to rely on language models to generate answers [4]. Constructive alignment ensures that the use of AI supports rather than undermines learning, by designing assessments that require application, reflection, and critical engagement, rather than mere information recall.

Another foundational concept is the distinction between declarative knowledge (knowing what) and functioning knowledge (knowing how) [22]. While AI tools can easily supply declarative knowledge, students must develop functioning knowledge through experience, practice, and feedback [11]. Creating the conditions for such learning requires setting the stage for effective teaching. Creating conditions means clearly defining the learning outcomes from the outset, understanding the students' prior experiences, and selecting teaching methods that actively engage them in the learning process [24]. It also means recognizing the contextual factors that shape education—class size, institutional culture, technological infrastructure, and now, increasingly, the availability of AI tools. Teaching is always situated within these broader contexts, and educational design must be responsive to them [50].

Ultimately, many pieces of literature emphasize the importance of student-centered learning environments. These are environments where students are empowered to take ownership of their learning and where teaching is designed to meet learners where they are [50][22][18]. In a world where students increasingly turn to AI tools for support, student-centered learning becomes even more critical. Rather than banning or ignoring these tools, educators can design learning experiences that teach students how to use AI responsibly and reflectively, encouraging Susan-like engagement rather than reinforcing Robert's avoidance strategies.

6.7 Assessment of student learning

Measuring student learning is a critical aspect of higher education, serving multiple purposes: it tracks student progress, evaluates instructional effectiveness, and supports institutional accountability and quality assurance. Given this research's focus on the impact of LLMs on education and the adjustments required in response to this shift, it is important to examine the existing methods used to assess student learning globally. These include traditional assessment methods, such as exams, grades, and standardized tests, as well as formative and summative evaluation practices (see Table 6.8). Additionally, authentic assessment models, including portfolios, project-based work, and capstone experiences, are utilized to measure real-world competence (see Table 6.1). Maintaining rigorous and reliable assessment practices remains central to ensuring that higher education institutions produce skilled and capable graduates [51].

The fundamental objective of any assessment is to determine how much a student has learned following a course. All measurement techniques, whether tests, assignments, or alternative assessments, are ultimately tools intended to support this goal. However, this paper does not focus on evaluating the effectiveness of existing assessment methods, despite ongoing debates in the literature about their limitations. Numerous studies have questioned both the pedagogical value and the accuracy of these conventional approaches [52]. Nonetheless, examining those critiques in depth falls outside the scope of this research. Instead, we aim to understand how student learning is currently measured and to explore whether these existing methods remain effective in the wake of the emerging influence of LLMs in education.

Definitions and types of assessment

There are different methods for measuring student learning, ranging from traditional assessments to more contemporary approaches. Traditional methods include exams, grades, and standardized tests, which offer a data-driven means of evaluating learning outcomes (see Table 6.1). Standardized tests, in particular, aim for broad content validity by focusing on transferable skills rather than course-specific knowledge. The most common measures of student learning remain course-based assessments such as exams and assignments. These include midterm and final exams, lab reports, essays, and other graded tasks that serve as core indicators of student performance. Their theoretical foundation is rooted in classroom pedagogy and curriculum design: in a well-aligned course, assessments are designed to reflect and measure the stated learning objectives [21]. However, in response to the limitations of traditional assessments, many educators in higher education advocate for the use of authentic assessment. This approach emphasizes the evaluation of students through meaningful, real-world tasks that demonstrate the application of knowledge and skills in practical contexts.

Quality criteria for assessment

To ensure that assessments accurately reflect the quality of student learning, specific quality criteria must be established and upheld. The literature is largely consistent in identifying four foundational pillars that underpin high-quality assessment: validity, reliability, usability, and transparency [53].

Validity refers to the extent to which an assessment measures what it is intended to measure. It ensures alignment between test content and intended learning outcomes, allowing for the drawing of meaningful conclusions from student performance. Without validity, the interpretation of results becomes speculative at best [53][54]. Reliability concerns the consistency of assessment outcomes across different conditions, such as multiple administrations or evaluators. Reliable assessments minimize random error and subjective bias, allowing for stable and replicable results [53][55]. Usability, also referred to as feasibility or practicality, addresses the extent to which assessments are manageable for both instructors and students. It involves considerations of time, resources, workload, and administrative efficiency. Assessments that are overly complex or time-consuming may undermine their effectiveness in practice [53][54]. Transparency ensures that students clearly understand how their performance will be evaluated. This evaluation includes explicit communication about grading criteria, the structure and purpose of assessments, and the feedback process. Transparent assessments foster fairness, increase student trust, and support informed learning strategies [53][55].

Overview of learning assessment [52][36][35][10]

Method	Theoretical Basis	Validity	Effectiveness
Grades	Grounded in classroom pedagogy and curriculum objectives; often teacher-designed. Includes final exams, quizzes, assignments contributing to a course grade. Implementation varies by instructor and discipline.	Reliability varies: objective exams can be quite reliable if well-constructed, whereas essay/project grades may vary between graders. Grades are a composite measure (content knowledge, skills, effort), so to fully interpretate, context is required.	Effective in certifying mastery of course content and immediate skills. Good at capturing knowledge and problem-solving within the taught context. May not capture ability to apply skills outside the course or long-term retention.
Standardized Tests	Uses common tasks/questions for all students. Often multiple-choice or performance tasks administered under uniform conditions.	High internal consistency and scoring objectivity yield strong reliability for group comparisons. Validity is strong for general skills if test content aligns with core outcomes, but may be limited to the specific constructs tested (e.g. analytic writing)	Effective for benchmarking core cognitive skills across programs/institutions. Captures learning in areas like critical thinking or quantitative reasoning in a standardized way. Less effective for creative, collaborative, or domain-specific skills not amenable to standardized items. Provides broad indicators of learning, but might miss nuance of individual student growth.
Formative Assessments	Based self-regulated learning theory. Implemented as low-stakes quizzes, drafts, discussions, etc., to give feedback throughout a course. Often ungraded or low-stakes, used by instructors and students to inform next steps.	Typically informal, so statistical reliability is not measured; instead, quality is judged by timeliness and clarity of feedback. Validity is usually high in context (activities target current learning goals), but results are not generalizable beyond the immediate cycle. They are not standardized – but that flexibility is by design.	Highly effective in improving learning processes and intermediate skills. Provides early evidence of learning or misconceptions, allowing interventions. Captures potential learning (what students can do with guidance) and builds toward summative outcomes.
Authentic Assessments	Draws on constructivist and situated learning theory; involves real-world or complex tasks. Implemented with open-ended tasks evaluated via rubrics and human judgment. Often used at end-of-course or program.	With clear rubrics and rater training, can achieve good reliability (faculty scorers can reach high agreement). Still, some subjectivity and task-specific variance. Tasks have high content and ecological validity. Scoring aligns with desired competencies.	Very effective at capturing complex learning, integration of knowledge, and higher-order skills (e.g. analysis, creativity, communication). Provides rich evidence of what students can do. Often promotes learning through the assessment itself.
Learning Analytics	Based on data science and learning science; continuously collects learner data (LMS logs, clickstream, interaction, etc.) to infer engagement and progression. Used via dashboards, predictive models, or adaptive systems. Can guide real-time teaching or support.	Data capture is automatic and consistent; interpretation reliability depends on robust models. Behavioral metrics may not consistently equate to learning. Validity depends on construct alignment; still evolving. Ethical concerns: privacy and bias must be addressed.	Effective for monitoring learning *process* and providing early warnings. Captures participation, study habits, and struggles that traditional tests may miss. Can show metacognitive behaviors. Best when used to inform interventions and support. Weak as a summative tool.
Developmental	Grounded in developmental psychology and college impact research. Involves tracking the same students over time to assess growth. Examples: entry-exit critical thinking tests, longitudinal surveys, interviews, portfolios. Often part of program evaluation or research.	Measuring change is difficult: gain scores have lower reliability, so large samples or multiple measures help. Practice effects and attrition must be managed. Validity is strong when instruments match developmental goals. External factors (internships, jobs) may confound results.	Crucial for capturing actual growth in cognitive or metacognitive abilities. Demonstrates value-added learning over time. Helps identify when learning occurs and informs curriculum. Not useful for short-term feedback but essential for program improvement and accountability.

Table 6.8: Overview of learning measurements, their theoretical bases, and their perceived validity and effectiveness, focusing on appliance and effectiveness

6.8 Deep dive into Visible Learning theory

Visible Learning, based on the well-received book by John Hattie, synthesizes and identifies key indicators of effective Learning [24]. It is widely regarded as an industry standard and a foundational text in the field of education. This work is relevant to this thesis, as it helps establish what constitutes effective educational methods and why they work. These insights serve as a basis for analyzing how LLMs might influence the effectiveness of existing practices. Furthermore, the findings may offer guidance on how to adapt or re-establish effective practices in light of the changes brought by GenAI tools such as LLMs.

The Visible Learning model is a research-based framework that synthesizes findings from over 1,600 meta-analyses. Its primary objective is to identify which factors have the most significant positive impact on student learning outcomes. A statistical measure called effect size (d) is used to rank influences on achievement. An effect size of 0.4 is called the "hinge point", representing a typical year's growth in Learning. Influences above 0.4 are considered high impact [24].

Before delving deeper into the theoretical underpinnings of Visible Learning, it is important to understand its core principles. Central to the findings is the notion that Learning becomes "visible" when teachers see learning through the eyes of their students, and students come to see themselves as their own teachers. The model emphasizes the importance of clarity in instruction and the central role of feedback in the learning process. Key high-impact practices identified include teacher clarity, formative assessment, effective feedback, and student self-assessment [24]. While this provides only a foundational overview, it establishes essential concepts needed to explore Hattie's theory in greater depth.

The Visible Learning model

The purpose of the Visible Learning model [24] is to improve student achievement by making the processes of teaching and learning clear, intentional, and evidence-based. It aims to ensure that teachers understand their impact on student learning and make informed decisions based on research findings. The Visible Learning model is based on five core perspectives: the why, how, what, doing, and evaluating.

1. **Why: Being clear about the purpose** – teachers see learning through the eyes of the students and encourage students to become their own teachers.
2. **How: How educators and students think** – highlighting the importance of mind-frames of students and teachers.
3. **What: Intentional alignment** among decisions about the knowing-that, knowing-how, and knowing-with aspects of the curricula, cognitive task analysis, optimal teaching interventions, and assessment or evaluation strategies.
4. **Doing: The implementation** based on the feedback loop of discover, design, deliver, double-back, and double-up.
5. **Evaluating: Evaluative thinking** – Parallel to all aspects of the model and guiding towards continuous improvement.

The mind frames speak to the why of teaching and learning, an intentional alignment model. Mind frames are the belief systems that drive intentional, reflective, and impact-focused teaching. Adopting and discussing these mind-frames is crucial for creating visible learning environments where both students and teachers are active, aware, and responsible for learning progress [24].

Knowing-that, knowing-how, and knowing-with represent three levels of understanding. Knowing-that refers to factual or conceptual knowledge, knowing-how involves the ability to apply skills and processes, and knowing-with is the capacity to integrate and use knowledge flexibly in real-world contexts. Together, these forms of knowledge support deeper learning by moving students from basic recall to meaningful application and transfer [22].

Double-back and double-up are strategies to ensure all students make meaningful progress. Double-back involves revisiting and reteaching content when students have not yet achieved the intended learning outcomes, ensuring no one is left behind. Double-up, on the other hand, is about extending and deepening learning for students who are ready for more challenging tasks, offering enrichment or more complex activities. Together, these approaches support differentiated instruction and promote both equity and excellence in the classroom [22].

The Intentional Alignment model refers to a cohesive connection between curriculum (what is taught), pedagogy (how it is taught), and assessment (how learning is measured) [24].

1. 1. Define clear learning goals and success criteria based on curriculum expectations and students' past, current, and future learning needs.
2. 2. Break down the lesson into cognitive tasks related to factual knowledge, skills, and knowledge transfer, ensuring students understand the complexity and what success looks like.
3. 3. Create a supportive classroom culture where mistakes are viewed as learning opportunities and where all students feel safe and included.
4. 4. Match teaching strategies to the task's difficulty and support students in building the confidence and ability to meet their learning goals.
5. 5. Provide students with effective strategies to manage and succeed in challenging learning tasks.
6. 6. Select activities that correspond to the depth of the content and encourage higher-order thinking and conceptual understanding.
7. 7. Use assessments that measure factual understanding, focusing on the content (knowing-that), the relational (knowing-how), and transfer (knowing-with) aspects of the success criteria

Implementing these seven steps involves the five steps from the 5D model: discover, design, deliver, double-back, and double-up. This model is used to help frame knowledge about the quality of implementation.

Summary of the 5D Model [24]

Phases	Teachers	Students	Families	Schools
D1: Discover Identifying goals that are worth pursuing above all else, building a theory of the pre-sent, and agreeing what success looks like.	Developing a deep understanding of their students' backgrounds and prior learning experiences.	Bring with them a rich set of values, expectations, and prior learnings rooted in their own context.	Act as their children's first teachers, shaping early attitudes and understandings about learning and success.	Called to support both academic progress and personal well-being, fostering a sense of unity and collective purpose.
D2: Design Systematically examining the different options in design space, selecting/designing a high-probability intervention, stress testing before launch. and developing your monitoring and evaluation plan.	Focus on aligning instruction intentionally by using methods such as backward design and thorough lesson planning. They address all levels of learning—surface, deep, and transfer—and conduct cognitive task analyses to match strategies with learning demands	Expected to understand learning intentions and success criteria, and engage meaningfully with feedback.	Support the design process by learning and using the language of learning, becoming attentive and active listeners, and maintaining high expectations for their children. They help create an environment where learning from failure is encouraged and recognized as a valuable part of the educational process.	Responsible for fostering a culture that supports learning through shared language and understanding of how learning happens. They must be aware of financial factors, sustain a climate of openness and invitation, maintain high expectations, and promote collective efficacy across the school community.
D3: Deliver Putting your agreed interventions into action and collecting monitoring and evaluation data.	Implement planned interventions with clear intent and purpose. A classroom culture that supports evaluative thinking and use evidence to inform their teaching. A variety of teaching methods are employed to meet diverse needs, and feedback—both given and received—is central to this process.	Play an active role by engaging with multiple learning strategies suited to different tasks and challenges. They are encouraged to invest effort and take ownership of their learning journey. Mistakes are seen not as setbacks but as valuable opportunities to grow. Collaboration is also key, with students learning to work effectively with others in pursuit of shared goals.	No specific implementation for families.	Support this phase by promoting a shared narrative that aligns everyone around common learning goals.
D4: Double-back Monitoring and evaluating your delivery chain and deciding where to go next.	Reflect on the impact of their teaching by monitoring the effectiveness of their delivery chain. This includes reteaching when necessary, critically evaluating their own impact and that of others, and helping students consolidate their learning.	Students are active participants in this phase by engaging with feedback, using self-assessment tools, and reflecting on their own learning progress. They work on assimilating new information and adjusting their understanding, demonstrating flexibility and growth in response to what they've learned.	No specific implementation for families.	"Trust but verify" approach—encouraging professional autonomy while maintaining accountability.
D5: Double-up either implementing and maintaining enhanced versions of an intervention in the same local setting and/or working to embed it across multiple schools.	Focus on sustaining and expanding successful interventions. This includes investing in their own professional development and enhancing their expertise. A critical skill at this stage is de-implementation—knowing when and how to stop practices that are no longer effective in order to make room for more impactful strategies.	Demonstrate a desire to go deeper in their learning. They show readiness and motivation to invest more time and effort into acquiring richer, more complex knowledge and skills. Their commitment to deeper learning signifies maturity and ownership of their educational journey.	Play a role by supporting extended educational engagement. They value continued learning and express willingness to support their children in pursuing more years of meaningful schooling, whether through formal education or sustained academic involvement.	Leaders help scale and embed effective practices by building a collective approach. They focus on upscaling interventions across schools or systems and, like teachers, must also develop the capacity for de-implementation—strategically discontinuing what doesn't serve the greater goal of sustainable improvement.

Table 6.9: Overview of the 5D model, designed to help implement the Intentional Alignment model

Overview of key findings

The literature on this topic is extensive; however, this section focuses only on the aspects most relevant to our research questions. To address these questions effectively, it is essential to highlight several key findings, which are outlined below.

There should be a focus on Learning Impact. Effective education involves teachers seeing learning through the eyes of their students and evaluating their impact, rather than fixating on particular teaching methods. The emphasis is on what students learn and using evidence (such as assessments and student feedback) to guide their practice. Every student is seen as capable of progress, and the teacher's role is to help each student exceed their current potential [22].

Feedback and Assessment are crucial: The process of making learning "visible" involves continual feedback loops – teachers gathering evidence of Learning and students receiving feedback on their progress. Feedback to students, when provided in a timely and specific manner, substantially improves their achievement. Likewise, teachers who regularly seek feedback on the effectiveness of their teaching (through assessments or student input) can adjust their methods to meet student needs better [22].

Thoughtful Use of Technology: Despite the proliferation of educational technology, Visible Learning synthesis finds that technology, by itself, has a relatively low average impact on Learning. Note that this was written before the rise of significant LLMs systems. This low impact is because many schools use tech as a direct substitute for traditional tools without fundamentally changing learning experiences. Visible Learning identified that technology is most powerful when used to enhance peer learning and provide multiple practice opportunities. Additionally, some students voice their thoughts and questions more openly via online tools or social media than in face-to-face classes, which can provide teachers with valuable insight into student understanding. The principle is that tech should augment, not just digitize, good teaching; for instance, enabling collaboration, personalization, and instant feedback, rather than replacing effective teacher-student interaction [24].

Certain factors harm learning: Chronic boredom, over-reliance on teachers or technology (excessive teacher control without student autonomy), and corporal punishment all showed negative effects on achievement. These underscore the importance of engaging lessons, fostering student agency, and maintaining a safe, non-threatening environment. Besides this, there should be high Expectations for all. A universal principle is that teachers must hold high expectations for every student. Labeling students negatively can inadvertently lower expectations and become self-fulfilling [24].

Top influences on student achievement [24]

Influence	Effect Size (d)	Description
Collective Teacher Efficacy	1̃.30–1.34	The shared belief of a school's teachers that *together* they can positively influence students. This has the largest impact of any factor – schools with high collective efficacy can yield over three years of student progress in one year. It reflects a culture of collaboration and trust among teachers focused on student learning.
Self-Reported Student Expectations (Students' Predicted Grades)	1̃.30	How accurately students predict their own performance. This surprisingly strong effect reflects **student metacognition** – students who understand their learning progress can set goals and self-regulate. It's often seen as an *outcome* of effective learning strategies; teachers can leverage this by helping students reflect on their own learning.
Cognitive Task Analysis	1̃.29	Explicitly teaching and modeling **how to think through tasks and problems** (breaking tasks into steps, analyzing how to solve them). By revealing the cognitive processes of experts, teachers help students develop problem-solving strategies. This approach shows very high impact as it builds students' ability to tackle new challenges independently.
Response to Intervention (RTI)	1̃.29	A structured, tiered approach to early intervention when students struggle. RTI involves frequent progress monitoring and targeted teaching strategies to address learning gaps. The high effect size indicates that **providing timely, appropriate interventions** (especially in early grades or as soon as difficulties arise) significantly boosts achievement.
Piagetian Programs	1̃.28	Teaching strategies built on **Piaget's developmental stages** – i.e. tailoring learning experiences to students' cognitive development level and encouraging discovery learning appropriate to that stage. These programs (often hands-on and inquiry-based) have a strong impact by aligning instruction with how students learn at different ages.
Jigsaw Method (Cooperative Learning)	1̃.20	A collaborative learning strategy where students become "experts" in one part of a topic and then teach it to their peers (like pieces of a jigsaw puzzle). Jigsaw method combines individual accountability with group interaction. It yields large learning gains because **students teach each other**, engage actively, and practice communication.
Conceptual Change Programs	0̃.99	Instruction explicitly designed to **confront and replace misconceptions**. These programs challenge students' prior conceptions (especially in science/math) so that students reconstruct their understanding correctly. The near-1.0 effect shows the power of directly addressing misunderstandings in learning new content.
Classroom Discussion	0̃.82	Purposeful, teacher-guided **discussions of topics** in class. When students discuss, argue, explain, and question in a structured way, it deepens understanding and critical thinking. Hattie's synthesis finds that rich classroom dialogue – as opposed to lecture-only – has a very positive effect. It helps students clarify their ideas and learn from peers.
Interventions for Learning Needs	0̃.77	School-wide or programmatic interventions aimed at students with learning difficulties (e.g. special education programs, intensive tutoring). When done well (early identification and support), these interventions can substantially accelerate learning for struggling students. This underscores the importance of *targeted support*.
Teacher Clarity	0̃.75	The teacher's ability to communicate **clear explanations, goals, and expectations** in lessons. Clear instruction (breaking down content, setting objectives, checking for understanding) leads to better student outcomes. Students are less likely to be confused and more likely to stay focused on the learning intention.
Feedback	0̃.70	**Feedback** refers to information given to students about their performance and how to improve. Hattie found feedback (especially when it addresses the task and how to do better, not just a score) is one of the most powerful drivers of learning. It helps students correct errors and reinforces what they're doing well.

Table 6.10: Overview of the top influences on student achievement, updated by Hattie.

Implications for teachers

At the core of Visible learing is the concept of adopting a mindframe of impact. Teachers should view themselves not merely as content deliverers but as evaluators of their own teaching. This shift involves a continual process of inquiry: “"Is my teaching effective? How do I know?"”. To answer these questions, educators can utilize formative assessments, exit tickets, and student reflections, enabling them to make learning visible and respond accordingly when students are not making progress. This reflective, evidence-based approach underpins sustained professional growth and student success [24].

Clear learning goals are another cornerstone of effective practice. A course should begin with explicit learning intentions and success criteria, presented in language that students understand. This clarity not only sharpens students’ focus and enhances the relevance of activities but also empowers them to self-regulate by assessing their own understanding and determining appropriate next steps [24].

Feedback, when implemented purposefully, becomes a potent tool for improvement. Teachers are encouraged to provide timely, specific, and actionable feedback focused on how students can improve their work. Equally important is inviting feedback from students regarding the effectiveness of teaching methods. Feedback should emphasize the task and the processes for progress rather than personal evaluations. Teaching students how to interpret and act on feedback further maximizes its impact on learning [24].

The research highlights a range of high-impact strategies that teachers can incorporate into their practice. Structured classroom discussions and cooperative learning techniques, such as the jigsaw method, engage students deeply and promote understanding through peer interaction. Explicit instruction in study skills—like summarizing, memorizing, and using mnemonics—enhances content retention. Regular formative assessments help consolidate prior knowledge and guide new learning. Moreover, peer tutoring and feedback activities leverage the social dimensions of learning and further reinforce comprehension [24].

Balancing knowledge acquisition with the development of higher-order thinking is another critical instructional priority. Teachers are urged to integrate direct instruction of foundational knowledge with inquiry-based activities that foster critical and creative thinking. This intentional alignment ensures that students not only master the content but also learn to apply and extend it in a meaningful way. In this dual approach, knowledge and skills are mutually reinforcing, creating a richer educational experience [24].

Ultimately, fostering a classroom climate that encourages risk-taking and views mistakes as integral to the learning process is crucial. Teachers can normalize error-making by modeling it themselves, reinforcing the message that struggle is a natural and valuable component of learning. Strong teacher-student relationships, built on trust, care, and mutual respect, form the foundation of such an environment. Although not the highest-impact factor in isolation, these relationships enable the successful implementation of other high-effect strategies [24].

6.9 Impact of Artificial Intelligence on education

What is the relationship between the capabilities of current LLM technology and educational theories on effective learning and teaching? This relationship is a crucial aspect of the central question addressed in this paper. Modern LLMs can automate various types of assignments and assessments, in some cases rendering traditional formats obsolete. In other contexts, LLMs can serve as personalized assistants, extending instructional methods and contributing to richer, more engaging learning experiences.[28].

Impact and validity of Artificial Intelligence in measuring student learning

LLMs impact traditional educational measurement tools, as they now can significantly influence the outcomes of these assessments. As a result, not all conventional measurement methods remain valid or reliable. In some cases, assessment practices need to be adapted. In others, they have become entirely obsolete [3].

There are various ways to assess student learning outcomes during or at the end of a course. These can include tests, assignments, graded group projects, and more [22]. A comprehensive overview is provided in Table 6.8. Today, advanced LLM platforms are capable of performing many of these assessment tasks, often at a level that surpasses the average student [27]. These capabilities raise concerns about the integrity of traditional assessment methods, many of which may no longer be valid or reliable. Take-home exams, reports, homework, (coding) assignments, and group projects can now often be completed within minutes or hours using publicly available LLMs [27]. While LLMs are also capable of handling test-style questions, their impact largely depends on how strictly students are monitored and restricted in their use of such technologies during examinations.

Relation between Visible Learning, effective education, and Artificial Intelligence

The purpose of the Visible Learning model is to improve student achievement, based on five core perspectives: the why, how, what, the doing, and the evaluating [22]. Specifically, the "Doing" and "Evaluation" parts are affected by current Popular LLMs.

The "evaluation" step is impacted since traditional ways of measuring can be compromised by LLM [27], making it harder to determine how students are performing and what they have learned, and creating feedback loops to establish a positive learning cycle. The "doing" step is affected since many of the essential parts of learning, that is, a feedback loop of discovery, design, and delivery, could be automatically done by LLMs, completing the task without the student learning anything. An essential part of learning is practicing and applying acquired knowledge to develop a deep understanding [42]. When LLMs automate this step, the opportunity for meaningful learning is reduced.

Benefits of Large Language Model technology for education

LLM technology can offer several benefits to students' learning. It can provide personalized support by offering tailored explanations, feedback, and guidance based on individual learning needs [3]. LLMs can also enhance language and writing skills by improving grammar, structure, and clarity [27]. They can aid in idea generation and brainstorming for essays, projects, and research, helping students overcome creative blocks [4]. Additionally, LLMs support practice and review by generating quizzes, practice problems, and step-by-step explanations that reinforce understanding [4]. In technical fields, LLMs assist with coding tasks, including debugging and concept clarification. They also help students manage study materials more effectively by summarizing content and rephrasing complex information. Overall, LLMs promote greater learning autonomy, allowing students to explore topics at their own pace and take more control over their educational progress [17].

6.10 Current developments in the landscape of Artificial Intelligence in education

The widespread availability of highly capable LLMs began only in 2023, just two years prior to the writing of this paper. Although some research has been published, it is still emerging gradually in response to the rapid rise of these systems. High-quality academic work requires time to develop, especially in a field that is evolving this quickly. Meanwhile, LLMs continue to improve significantly, with their capabilities advancing fast since their first release. Educational institutions and governments respond much faster to this technological shift. For this reason, this thesis also examines current developments within the educational landscape, including how institutions and governments are responding, what experts are saying, and how students perceive these changes. While not all of this information is strictly based on peer-reviewed research, it still offers valuable insights into the broader impact of LLMs on education.

UNESCO's framework on Artificial Intelligence and education: Guidance for policy-makers

As AI continues to reshape the social, economic, and educational fabric of societies worldwide, UNESCO has responded with the development of two targeted AI competency frameworks: one for students and one for teachers. These frameworks provide structured guidance to support education systems in conveying the knowledge, skills, and ethical understanding necessary. UNESCO distinguishes AI from other digital technologies due to its profound societal impact and the complex ethical challenges it presents, such as bias and accountability. The frameworks reflect this distinction by addressing competencies beyond traditional digital literacy [25].

The AI competency framework for students aims to equip young learners with the capacity to engage thoughtfully with AI technologies. It includes four core areas: developing a human-centered mindset, fostering an understanding of AI ethics, building foundational knowledge of AI techniques and applications, and promoting creative engagement through AI system design. These components are designed to be integrated throughout the curriculum, fostering interdisciplinary learning that spans both technical and social domains. The framework is not merely about using AI tools but about cultivating an informed and critical perspective on their design and use [56].

For teachers, the competency framework emphasizes professional growth and responsible integration of AI in pedagogy. It is organized into five domains: promoting a human-centered mindset, understanding and teaching the ethics of AI, acquiring foundational AI knowledge, applying AI in innovative teaching strategies, and using AI tools for personal professional development. Central to this approach is the principle that AI should enhance rather than replace the human dimensions of teaching. The framework positions teachers not only as facilitators of AI literacy but also as ethical stewards guiding students through increasingly complex digital environments [57].

UNESCO's human-centered approach is reflected in both frameworks. Drawing from its broader policy foundations, including the 2021 Recommendation on the Ethics of AI and the 2019 Beijing Consensus, UNESCO emphasizes that AI must be used in ways that uphold human dignity, advance social justice, and respect cultural diversity. AI is seen as a tool to support human development, not to supplant it [25]. Beyond curriculum design, the frameworks carry broader policy implications. UNESCO advises that countries adopt a comprehensive strategy for AI capacity building across all levels of education. This includes ensuring equitable access to digital infrastructure, promoting environmentally responsible AI practices, and avoiding an over-reliance on AI to address structural issues such as teacher shortages or underfunded schools. Instead, sustained investment and policy reform are needed to address these foundational challenges [25].

OECD on Artificial Intelligence and education

The The Organisation for Economic Co-operation and Development (OECD) publishes reports on the future of education, economic opportunities, technological innovations, and more. Understanding the societal impact of these developments, especially within the education system that prepares students, requires a clear comprehension of AI's capabilities and its developmental trajectory. Furthermore, comparing AI's abilities with human skills is essential to determine where AI can replace human actions and where it can complement them. The OECD has also concluded that AI, including LLM systems, now surpasses humans in nearly every education-related task, prompting educational institutions to rethink current practices, supporting the very purpose of this thesis. The report emphasizes the importance of acknowledging AI's current capabilities and carefully considering their implications for the future of education [37].

"... determine which skills to prioritize, which to phase out, and where to place greater emphasis in an AI-influenced world..."[37].

Similar to the focus of this thesis, the OECD emphasizes the importance of anticipating how learning methods and teaching practices will evolve. It also emphasizes the importance of clearly defined educational goals for future generations. Further research should be encouraged on the effective integration of GenAI in teaching and learning processes [58].

"... tasks humans perform today are likely to change in the future. In consequence, exploring the way humans use, rely on, or collaborate with AI's is necessary to adjust the way we should rethink education systems in light of AI capabilities ..."[37].

The education system should not only adapt to the rise of AI in order to remain as effective as it was before, but also respond to how AI will reshape the world and, indirectly, the current way jobs are done. The impact of AI on education extends beyond questions of effectiveness; it also involves shifts in goals, approaches, and societal values [37].

European School Education Platform on Artificial Intelligence for teaching and learning

The European School Education Platform (ESEP), the European Union's central hub for school-level educational resources and collaboration, has taken an active approach to integrating AI into education. At the foundation of ESEP's position is the belief that AI and data literacy are essential components of digital competence in the current data-driven societies. The platform emphasizes that educators must not only be able to use AI tools but also critically assess AI-generated educational content. It urges schools to consider both the pedagogical value and the ethical dimensions of AI, such as data privacy, algorithmic bias, and the role of human agency in automated processes [59].

ESEP promotes a human-centred, ethically guided approach to AI in education, aligning with broader international guidance such as that of UNESCO [25] as Petra Bevek from the Slovenian Ministry of Education noted that there is a pressing need for international dialogue about the risks and benefits of AI, particularly in the context of teacher education [59]. ESEP views the preparation of teachers and school leaders as a strategic priority. Initiatives like the AI4T (AI for Teachers) project provide online courses and multilingual open textbooks to equip educators with the foundational knowledge and confidence needed to use AI effectively and responsibly.

The platform also highlights practical examples of AI implementation across Europe. One notable case is the Lycée des Arts et Métiers in Luxembourg, where educators use AI tools to facilitate student creativity, personalize learning, and reduce the technical barriers of complex software [60]. Teachers act as coaches, supporting students as they engage with GenAI tools such as ChatGPT, image generators, and AI-assisted design software. This approach aligns with ESEP's broader endorsement of blended learning, adaptive teaching, and learner autonomy [61]. ESEP supports the introduction of AI concepts across all educational levels. At the secondary and vocational levels, projects such as AI4EDU, AIware, and AI4STEM are piloting curriculum-aligned AI assistants, lesson plans integrating AI with STEM and IoT principles, and even a certification framework for AI-ready schools [59].

Estonia sets the global standard for Artificial Intelligence in education

Estonia is launching a national initiative, AI Leap 2025, to integrate AI into its education system. Inspired by the successful Tiger Leap program of the 1990s, providing access to digital resources in education, AI Leap aims to equip students and teachers with essential AI skills and tools to maintain the country's competitiveness in the digital age. The program will begin in autumn 2025, offering 20,000 high school students and 3,000 teachers access to advanced AI-based learning applications, along with teacher training. It will later expand to vocational schools, ultimately reaching around 58,000 students and 5,000 teachers by 2027. Initiated by President Alar Karis and supported by both government and private-sector partners, including OpenAI and Anthropic, the program emphasizes equitable access, curriculum co-creation, and large-scale AI skill development. The long-term goals include improving educational quality, promoting digital inclusion, and preventing a AI literacy divide. Estonia is among the first countries to implement such a comprehensive, nationwide AI education strategy [62][63][64][65].

In the latest PISA assessment conducted in 2022, with results released in 2023, Estonia ranked highest in Europe in mathematics, science, and creative thinking, and came second to Ireland in reading. Building on this success, Estonia is launching a nationwide AI initiative designed to equip both students and teachers with advanced artificial intelligence tools and skills. [65] As part of the program, teachers will receive training focused on digital ethics, self-directed learning, and ensuring equitable access to AI education. According to officials, the initiative aspires to make Estonia not only one of the most digitally advanced countries but also a global leader in practical AI use in education [65]. Teachers will retain autonomy in deciding how to integrate AI into their teaching. Starting in September, students will receive personal AI accounts and be regularly encouraged to use their devices in the classroom.

Japan's Policy Approach to Artificial Intelligence in Primary and Secondary Education

Japan is taking a proactive, structured, and human-centered approach to integrating GenAI into its education system. The Ministry of Education, Culture, Sports, Science, and Technology (MEXT) has developed a comprehensive framework to support educational institutions in navigating this transition. They have been researching this since 2023 and continuously [66].

MEXT published the Guidelines for the use of GenAI in Primary and Secondary Education (Version 2.0) in December 2024 [67]. These guidelines reflect updated insights and build upon the interim version first released in July 2023. Rather than imposing strict regulations, the guidelines provide flexible, scenario-based recommendations to encourage responsible and context-aware use of AI in educational settings [67].

The core philosophy is grounded in three principles [67]:

- **Human-Centered AI Use** – AI should serve as a tool to extend human capabilities, not replace them.
- **Purpose-Driven Integration** – AI should be used only when it supports educational objectives.
- **Ethical and Literate Engagement** – Teachers and students alike must understand the risks of AI, including bias and misinformation, and use it responsibly.

These principles are aligned with Japan's broader national AI policies, including the Human-Centered AI Social Principles and the Hiroshima AI Process, which emphasize dignity, fairness, and sustainability. MEXT distinguishes between the use of GenAI by educators and students, offering detailed guidance and case examples: [67]

Teachers are encouraged to use GenAI to:

- Prepare lessons, quizzes, and administrative documents.
- Simulate student responses for teaching preparation.
- Summarize meeting notes or lecture recordings.

Teachers must review and revise AI-generated outputs to ensure accuracy and reliability. Usage must comply with Japan's data protection laws and copyright regulations, particularly when AI outputs are disseminated beyond the classroom or repurposed for public use.

Student use of AI is promoted when it contributes to [67]:

- Deepening understanding through alternative explanations or perspectives.
- Enhancing creativity, especially in group discussions and writing processes.
- Supporting language learning and inclusive education.

Overall, inappropriate uses are also identified [67]:

- Submitting AI-generated content for assessments or contests as original work.
- Replacing core expressive or reflective learning tasks (poetry, personal essays).
- Relying solely on AI output for knowledge without critical engagement.

MEXT emphasizes that AI should not undermine the purpose of learning. Tasks must be restructured where necessary to maintain academic integrity and encourage meaningful engagement.

To support practical implementation, MEXT has established a AI Pilot School initiative. Over 50 schools across Japan have been designated to experiment with AI integration in both classroom instruction and school administration. These schools serve as testing grounds for developing effective practices, addressing challenges, and providing feedback to refine national guidance [68][67].

Japan's transition to AI-supported education is marked by a policy stance centered on guidance rather than enforcement, granting schools the autonomy to adopt AI in ways that align with developmental and educational goals [66]. Alongside this, capacity building through teacher training and institutional support is emphasized to promote responsible implementation. Policies are subject to continuous review and adjustment in response to evolving technology and classroom feedback. This structured yet flexible strategy reflects Japan's commitment to integrating AI into education in a way that balances innovation with pedagogical integrity and ethical responsibility [66].

China's adaptive learning platform that personalizes K-12 education

Squirrel AI is an example of how LLM technology is being integrated into educational technologies to support adaptive, personalized learning at scale. Founded in 2014 in Shanghai, Squirrel AI positions itself at the forefront of AI-powered tutoring systems. The company's core product is its Intelligent Adaptive Learning System (IALS), which uses a fine-grained model of knowledge mastery to individualize instruction for each learner. IALS decomposes academic content, particularly in mathematics and language arts, into thousands of "knowledge points." These are continuously assessed and remapped based on a learner's interactions, enabling a personalized learning path that adapts in real time to student needs [69].

In early 2024, Squirrel AI introduced its Large Adaptive Model (LAM), an advanced system built on a transformer-based architecture. The LAM leverages data from over 24 million students and more than 10 billion learning behavior records to enhance the precision of its predictions and interventions [70]. This model not only improves the targeting of instructional content but also supports modules for emotional engagement, motivational scaffolding, and real-time feedback. For example, LLM-based components are employed to provide dynamic error analysis, automated essay feedback, and even empathetic tutor-like responses. The result is a hybrid instructional approach that blends adaptive algorithmic decision-making with natural language interactivity [71].

The company has rapidly expanded across China and into international markets, with over 3,000 learning centers operating in more than 200 cities. Its partnerships include collaborations with research institutions such as Carnegie Mellon University, the University of California, Berkeley, and SRI International. The U.S. division of Squirrel AI, led by Dr. Joleen Liang, focuses on LLM research and development, particularly through its dual headquarters in Seattle and Shanghai [69].

Squirrel AI exemplifies the current trajectory of LLM deployment in education, where natural language capabilities are not limited to content generation but are embedded into intelligent learning systems that drive assessment, feedback, and motivation. Its success highlights a shift from general-purpose educational software toward highly adaptive, data-intensive platforms capable of simulating individualized human tutoring. As such, it serves as a compelling model for how LLMs can be operationalized in traditional and emerging education systems, aligning closely with global discussions on the responsible and effective integration of GenAI in learning environments.

United Kingdom's solution to support mathematical education

Eedi is a UK-based educational technology company focused on improving mathematics understanding for students aged 9 to 16 through diagnostic assessment, adaptive learning, and on-demand tutoring. Eedi has gained traction in schools across the UK and internationally, with over 160,000 teachers and 19,000 schools reportedly using the platform. Eedi's primary innovation lies in its ability to identify and address student misconceptions in real time using AI-powered diagnostic tools and learning pathways [72].

At its core, Eedi delivers diagnostic quizzes that help teachers pinpoint specific conceptual gaps among students. These questions are carefully designed to expose common mathematical misconceptions—for example, misunderstandings about place value, negative numbers, or ratio reasoning. Once a misconception is identified, the system delivers targeted follow-up materials such as short video lessons, hints, or scaffolded tasks, many of which are guided by AI-based decision systems. In premium versions of the platform (Eedi Plus), students also gain access to qualified human tutors via a live chat interface [72].

AI is embedded throughout Eedi's pedagogical model. The platform uses LLMs to generate and refine diagnostic questions and explanations, ensuring that each query is accessible, unambiguous, and aligned with curriculum standards. It also employs adaptive algorithms to select the most appropriate next question based on a student's response history—an approach designed in collaboration with Microsoft Research [73]. Eedi's approach reflects a "human-in-the-loop" model of AI in education, where AI-generated outputs are consistently reviewed or augmented by expert teachers. This balances the efficiency of algorithmic decision-making with the contextual judgment of professional educators. The platform's commitment to pedagogical soundness has led to multiple trials and evaluations by organizations such as the UK's Education Endowment Foundation, which has found preliminary evidence of positive effects on student outcomes in math when Eedi is implemented effectively [44].

Singapore's Artificial Intelligence strategies AI4E and SLS

AI for Everyone (AI4E) is a initiative by AI Singapore, intended to build foundational AI literacy among the general public. Since its inception, the programme has been a core component of Singapore's national strategy to equip citizens with the knowledge and mindset needed to thrive in an AI-enhanced world. It aligns closely with the country's Smart Nation vision and the Ministry of Education's (MOE) broader EdTech Masterplan 2030, which aims to transform education through technology to prepare students for a digitally advanced future [74]. As of 2025, AI4E has been iteratively updated, with version 4.0 comprising six modules: What is AI?, AI is not magic, What can AI do?, GenAI, Responsible AI, and AI, Jobs & You. These modules are offered in both online and face-to-face formats. Participants who complete the course receive a digital badge. AI4E serves learners of all ages—from students in secondary schools and ITEs to working professionals and senior citizens—demonstrating its broad inclusivity [75].

Singapore's broader push to integrate AI in education is also exemplified by its development of the Student Learning Space (SLS), which has undergone significant upgrades since its launch in 2018. Informed by the COVID-19 shift to home-based learning, and accelerated by the EdTech Masterplan 2030. ALS allows students to follow personalised learning paths in subjects like Mathematics and Geography, adapting to individual progress and mastery. ACP helps teachers generate AI-assisted lesson plans based on simple prompts, thereby saving time and allowing more focus on student engagement. ShortAnsFA and DAT offer AI-driven feedback and data analytics capabilities, giving teachers insight into student learning patterns and freeing them from repetitive tasks. The Appraiser Testimonial Generator is another example, streamlining the process of writing student testimonials using AI-driven drafts [74].

Prestigious universities' stance on Artificial Intelligence in education

Harvard University (North America) encourages the responsible and ethical use of GenAI tools across academic and administrative contexts. While acknowledging the potential benefits of AI for tasks such as brainstorming, drafting, and research [76], the university emphasizes the importance of transparency, privacy, and intellectual integrity. Users are advised to verify AI-generated content, avoid sharing sensitive information, and comply with course, departmental, and institutional policies. Overall, Harvard supports thoughtful experimentation with AI, provided it aligns with academic standards and ethical guidelines [31].

The Universidade de São Paulo (South America) has not yet issued a formal public policy on the use of AI in education. However, the university has established the Centro de AI (CIAAM) to conduct advanced research on AI applications, including those in education [77]. While proposals are emerging that explore AI 's potential to transform teaching, these ideas have sparked serious concerns regarding educational quality and equity. The ongoing discourse reflects both the promise and the tension surrounding AI 's integration into traditional educational settings [78].

ETH Zurich (Europe) advocates a proactive yet ethical integration of GenAI in education, guided by the principles of responsibility, transparency, and fairness. Both students and lecturers are encouraged to explore AI tools as creative and supportive aids—[79] but must disclose their use, verify outputs for accuracy and bias, and respect privacy and copyright rules [80]. Academic and legal standards are enforced through updated declarations of originality and disciplinary measures for misuse or non-disclosure of intellectual property. To support this, ETH provides workshops, library courses, and institutional funding to build AI literacy and integrate AI meaningfully into teaching, assessment, and scientific writing. Lecturers set clear course-level guidelines and model best practices, while students remain accountable for their work and must properly cite AI contributions. Through ongoing monitoring, training, and tools like Microsoft Copilot and dedicated workshops, ETH equips its community for a future in which AI enhances learning—without compromising academic integrity or human oversight [79].

Peking University (Asia) adopts a balanced and practical approach to GenAI, recognizing its growing role in legal education and practice. The school permits students to use AI tools for many learning-related tasks—such as summarizing cases, translating text, or brainstorming—so long as the same use would be acceptable if done by a person or conventional tool. However, by default, students may not submit AI-generated content directly in work evaluated for credit, even with attribution. Instructors retain discretion to modify these rules for their courses, provided changes are communicated in writing. STL emphasizes the importance of verifying AI output for accuracy and warns of risks related to plagiarism, hallucinated content, and privacy concerns. This policy encourages thoughtful use while upholding academic integrity and professional responsibility [34].

The University of Stellenbosch (Africa) has adopted a comprehensive and principle-driven approach to the ethical use of AI in both research and teaching-learning-assessment (TLA) [32]. Approved by the University Senate, its 2024 position statement outlines institutional values such as accountability, transparency, fairness, and authenticity as essential in guiding the use of AI. AI tools are viewed as supportive instruments that can enhance learning and research, particularly in contexts of linguistic diversity and resource inequality. However, their use requires clear disclosure, critical engagement, and human oversight [32].

Initiatives from Dutch schools and universities

There is a great deal of activity in the field of AIED, and research is only just keeping up. Educational institutions, such as universities, are launching initiatives, pilot programs, and conceptual projects to explore how AI is influencing education and how we can adapt to this digital transformation. Although these efforts do not constitute scientific data, they still offer valuable insights.

University of Amsterdam: UvA AI Chat

The UvA is developing the UvA AI Chat, a self-managed GenAI tool designed specifically for its academic community. One of the main concerns with using public AI tools is data integrity [81]. UvA AI Chat prioritizes privacy, data sovereignty, and ethical usage by keeping all user input within the university and offering customizable features aligned with educational and research needs. There are five main reasons for developing an in-house AI tool at the UvA:

1. Privacy and security
2. Control over data
3. Accessible and tailored
4. Sovereign and future-proof
5. Support for academic freedom

Currently in the pilot phase, and planning a first release after summer 2025, UvA AI Chat represents a strategic step toward institutionally governed AI in higher education [81].

TU Delft: Integrating Artificial Intelligence Across Education and Research

TU Delft takes a comprehensive, university-wide approach to embedding AI, data, and digitalization into both education and research. Through the TU Delft AI Initiative, the university supports curriculum innovation, interdisciplinary collaboration, and open sharing of educational materials to prepare students and staff for an AI-driven future [82][33].

To responsibly and effectively integrate AI into society, TU Delft emphasizes the need to educate both AI specialists (working in AI) and domain experts (working with AI). Specialists focus on the technical and ethical foundations of AI, while domain experts apply AI within their specific fields. TU Delft aims to ensure that every student receives education both in and with AI, fostering a future-ready, ethically grounded society [33].

> Everyone at TU Delft will have to deal with AI in one way or another. Therefore, every student should also be trained in AI, not incidentally, in a single subject, but where relevant as an integrated part of the curriculum.
> - Willem-Paul Brinkman, Academic Lead AI Initiative Onderwijs [33].

Among their initiatives are the AI Teachers Programme and the Open Educational Resources initiative. The AI Teachers Programme supports educators who teach in AI (technical and theoretical content) and with AI (domain-specific applications). It includes focus groups for curriculum integration and the Machine Learning Teachers Community, where instructors share experiences and best practices across faculties. The Open Educational Resources initiative promotes open, reusable AI learning materials that are accessible to students from diverse backgrounds and universities. A central repository is being developed in collaboration with other institutions, aimed at harmonizing standards across disciplines and universities [82].

TU Delft's research strategy combines fundamental and applied AI, aiming to keep all fields at the forefront of technological development. AI is approached not only as a technical challenge, but also through its interaction with ethics, governance, society, and human behavior [83].

AI innovation is organized around seven interdisciplinary domains:

- Human-centered, responsible AI systems
- Machine Learning
- AI for Energy & Sustainability
- AI for Health & Care
- AI for Ports & Maritime
- AI for Peace, Justice & Security
- AI for Advanced Industry

Radboud Universiteit & Rijksoverheid: NOLAI - A National Approach to implement Artificial Intelligence in Primary and Secondary Education

The Nationaal Onderwijslab AI (NOLAI) is the Netherlands' national initiative to responsibly AI into primary and secondary education. As the world's first publicly funded national AI lab focused on digital educational innovation, NOLAI is designed to support both pedagogical advancement and technological development. Led by Radboud University and funded through the Dutch National and European Growth Fund, the project began in 2022 and will run through 2032. It brings together educators, researchers, policymakers, and technology developers in a long-term, multi-stakeholder collaboration [84].

The project has two core objectives. First, it aims to develop intelligent, AI-based educational technologies that improve teaching quality and student outcomes in primary and secondary (including special) education. These objectives include innovations such as adaptive learning systems, automated feedback tools, and AI-powered support systems for teachers. Second, NOLAI seeks to understand better the pedagogical, ethical, and social implications of deploying AI in schools. Each technology developed under the program is co-designed with educators to ensure it is both effective and appropriate for real-world use [84].

> "We are investigating the operation of an adaptive system to reduce the need for children to take tests. The children are then less dependent on the result of one test."
> - Inge Molenaar, professor of Education and AI at Radboud University and general and scientific director of NOLAI [84].

NOLAI's impact is driven by its co-creation model. Throughout the program, it aims to complete 77 collaborative innovation projects with schools and research institutions. In addition, the initiative targets the development of 15 validated AI-driven educational tools, with at least 9 of these technologies to be demonstrated in classroom settings. The consortium also prioritizes knowledge sharing and capacity building, engaging with a vast network of schools and teacher-training programs to ensure sustainable adoption of AI innovations[84].

From an economic and societal perspective, NOLAI aims to enhance the quality of education nationwide. It seeks to enhance human capital, reduce teacher workload and shortages through AI support tools, and promote educational equity by enabling more personalized learning pathways. By addressing both technological and human factors, NOLAI serves as a national model for integrating AI into education in a responsible, evidence-based, and scalable manner [84].

Chapter 7

Discussion

In Chapter 6, an overview of the relevant literature and collected data was presented to address the research questions. This chapter builds on that foundation by providing answers to the research questions, drawing on the findings discussed earlier. Additionally, it highlights the study's limitations and outlines directions for future research.

7.1 Summary

The emergence of widely accessible and effective LLMs marks a technological turning point, comparable to the advent of the internet, the rise of the Google search engine, or the introduction of the iPhone [85] [86]. This shift is already transforming numerous sectors by automating human tasks and increasing efficiency across a wide range of activities [27]. Education is among the sectors significantly affected, requiring an urgent response within educational systems [4][3][2].

Educational institutions and related organizations have begun responding to this shift by exploring how to integrate AI effectively into teaching and learning [37][18]. Government initiatives, university policies, pilot programs, and emerging research all aim to address the challenges and opportunities presented by this technology [65][15][87][4]. These efforts seek to ensure that education remains effective, relevant, and aligned with the demands of a rapidly evolving digital landscape.

Learning and teaching are influenced by the capabilities of LLM-based tools [3]. Within the framework of the Visible Learning model [24], the "Doing" and "Evaluating" phases are particularly affected. Automation of key learning activities can diminish the effectiveness of these phases by reducing active student engagement [3]. For example, tasks such as essay writing can now be fully automated, allowing students to submit assignments without actively participating in the learning process. Similarly, in the evaluation phase, it becomes challenging to distinguish between a student's work and content generated by AI tools [2]. As a result, core course components, such as assignments, projects, portfolios, written reports, and research activities, are increasingly vulnerable to automation, raising critical questions about authenticity and the integrity of learning outcomes.

In response, the education sector must take deliberate action to ensure that learning remains effective and meaningful [12]. This response necessitates a reevaluation of traditional course structures and the development of new models that take into account the capabilities of AI. Insights from pilot programs, institutional initiatives, and emerging literature emphasize the importance of returning to the foundational principles that make education effective. The challenge lies in re-imagining educational practices in a way that preserves these core ideals—such as critical thinking, active engagement, and authentic learning [24] while integrating new technologies responsibly and constructively [17]. In the short term, curriculum design must be adapted to preserve academic integrity and maintain learning effectiveness in the face of rapidly advancing technology [82]. In the long term, this shift presents an opportunity to rethink education more broadly, focusing on how it can become not only more effective but also more engaging, thereby triggering the intrinsic motivation of students to learn.

7.2 What characteristics make traditional course settings effective as an educational method?

Traditional course settings have long been considered effective due to several key characteristics that support structured and purposeful learning [24]. First, they offer a clear instructional framework, typically grounded in well-defined learning objectives, sequential content delivery, and consistent assessment practices [50]. This structure helps guide learners through increasingly complex material while enabling instructors to monitor progress and provide timely feedback [88]. Second, traditional settings often promote active engagement through in-person interaction, classroom discussion, collaborative learning, individual deep thinking, and problem-solving, all of which contribute to a deeper understanding and critical thinking [24]. Third, they foster a learning environment where motivation, discipline, and accountability are reinforced by routine and social presence [38]. Finally, traditional settings support formative assessment, testing, and direct feedback, allowing instructors to adjust teaching strategies based on students' needs [22]. Together, these features create a stable foundation for effective teaching and learning, remarkably when grounded in evidence-based pedagogical practices [24].

7.3 How are traditional course components affected by the emergence of Large Language Models?

The emergence of LLMs has an impact on traditional course components, particularly those that rely heavily on written output, independent research, and take-home assessments. AI tools can partially or fully generate assignments such as essays, reports, and project work [26], raising concerns about the integrity of assessment outcomes [18]. A key challenge is that teachers often cannot determine with certainty whether a student has used tools like ChatGPT, and even if they suspect it, they typically cannot provide definitive proof. This challenge significantly complicates the enforcement of academic honesty policies and undermines trust in traditional assessment formats.

Additionally, LLM-based tools reduce the cognitive load on students by automating parts of the (deep) thinking and writing process [89]. While this allows for faster task completion, it also risks diminishing deep learning, as cognitive effort is essential to processing information, constructing knowledge, and achieving long-term retention [24]. Reduced engagement in cognitively demanding activities may lead to a superficial understanding, particularly when students rely heavily on AI-generated content without actively participating in the learning process [50]. At the same time, LLMs can serve as digital tutors or study companions [64]. They can assist students by explaining concepts, answering questions, generating examples, and supporting the development of ideas. This can enhance individual learning by providing immediate, personalized feedback and scaffolding, especially for learners who may lack access to other forms of academic support [2].

Furthermore, the efficiency enabled by AI tools frees up time that students would otherwise spend on repetitive or mechanical tasks. This efficiency opens up space for more creative thinking, more profound exploration of topics, and playful engagement with learning [39][2]. When used thoughtfully, LLMs can support more flexible and student-centered learning environments [67]. However, realizing these benefits depends on how such tools are integrated into course design, assessment practices, and learning support systems.

Important note on cognitive offloading

AI tools are designed to offload tasks and processes, allowing users to work more efficiently and attempt things once out of reach. In education, however, this creates a critical tension: the very tasks AI can automate are often those students need to learn to build foundational skills. While many tasks may become obsolete, specific abilities remain essential for learners to internalize, especially early in their development. Visible Learning theory emphasizes the importance of Cognitive Task Analysis—explicitly teaching and modeling how to think through problems by breaking them into steps and analyzing solution strategies [24]. If AI tools entirely do this process, students risk missing a crucial stage of learning where they develop the mental frameworks necessary for independent problem-solving.

This distinction should guide how we integrate AI into learning environments. In some instances, it must be a conscious, pedagogical decision not to offload a task to AI. For example, in programming, students only truly learn by doing [90]. While AI tools can generate code or fix syntax errors, that alone does not teach problem-solving, logic, or a deeper understanding of the structures of software development. The key question becomes: "What should students still learn to do manually, and what can be safely offloaded to AI?"

Despite technological advancements in the last decades, we are still typing code line by line—a slow and inefficient process by today's standards. Looking forward, much of programming may become more automated and hopefully more intuitive. However, even in this scenario, a strong need will remain for individuals who understand the underlying code, architecture, and logic behind what the machine generates. AI can support programming by handling repetitive or surface-level tasks—like locating missing semicolons, writing boilerplate code, and translating solutions into code—but it cannot (yet) replace the full human capacity for abstraction, creativity, and complex problem-solving [91].

In the context of programming education, this suggests a dual approach: students should be expected to understand and practice the basics of coding, how basic code works, and also why it is important and even enjoyable. At the same time, students must train in solving (complex) problems and, to a lesser degree now than before, translate those into instructions a machine can understand. AI tools can offload translation solutions, identify errors, automate repetitive code, or implement existing solutions. That is the essence of programming: not typing code, but solving problems and translating those solutions to executable code. AI can be a new abstraction layer, similar to the ones we already have from bytecode, programming languages, to Excel, and now, AI.

In short, the guiding principle can be summarized as follows: offload what is repetitive and non-essential; preserve what is critical and meaningful. This balance will differ by discipline, but in all cases, it requires deliberate instructional design focused on learning outcomes rather than traditional task completion.

7.4 What strategies can educators adopt to ensure meaningful student learning in the presence of Large Language Model technology?

The rise of LLM technology in education demands a fundamental shift in how we design courses, assess learning, and engage students [18]. Instead of resisting or banning such tools, educators should instead focus on preserving traditional practices, adapting practices to preserve the core goal of education: helping students build meaningful, lasting knowledge that serves them beyond the classroom [92].

One critical strategy is to educate students explicitly on how to use LLMs effectively and ethically [33]. This strategy includes instruction on prompt engineering, critical evaluation of outputs, and understanding the limitations and incorrectness of these systems [93][83]. Students must recognize that LLMs are not complete sources of truth; they are predictive models that can generate convincing but also incorrect or fabricated responses [91]. Teaching students how to question and validate AI content is as important as teaching them to use the tools in the first place [66].

Educators need to shift their focus from traditional practices to the core purpose of education. The goal is not simply to complete assignments or pass exams but to ensure students grasp essential concepts, develop critical thinking skills, and build knowledge they can apply in real-world situations [24]. Teachers should view assessments as tools to support learning rather than as the goal itself [50]. Re-centering on learning objectives creates space to question whether legacy practices, such as essay writing, remain effective in a world where LLMs can easily perform such tasks. Additionally, research highlights the value of play as a powerful learning tool. Emerging technologies could enable teachers to integrate play more effectively into their strategies, fostering an engaging, natural, and enjoyable learning experience while supporting deeper understanding [41].

Consequently, there must be a willingness to abandon outdated methods when they no longer align with learning outcomes. For instance, if LLM technology can easily write an entire essay, and if students are likely to use it regardless of policies, then relying on essay writing in its traditional form becomes ineffective [16]. This ineffectiveness does not mean abandoning writing altogether, but rather rethinking how and why teachers use this method, shifting towards collaborative writing, in-class drafting, or analysis of AI-generated texts.

In addition to this, educators should update and align their teaching methods in cases where the impact on traditional educational practices is negatively impacting student learning [24]. This redesign can consist of the four important pillars:

- Clarify the learning goals: Define exactly what students should know or be able to do.
- Connect these goals to future usefulness: Help students understand how these skills or concepts will serve them in academic, professional, or personal contexts.
- Design assessments that support the goals: Choose or create assessment methods that reflect and reinforce the intended learning outcomes.
- Test understanding in authentic ways: Use assessments that measure knowledge and skills directly, such as oral exams, applied tasks, or knowledge checks, rather than solely relying on written submissions.

Underlying all these strategies is the need for a shift in educational mindset. Students should be encouraged to see learning as something they do for themselves, not for the teacher or the grade [22]. Completing assignments is not about compliance, but about preparation. If a student chooses not to engage in certain tasks, that may be acceptable, provided they can still demonstrate that they have achieved the learning outcomes. This mindset fosters autonomy and responsibility, shifting education toward a more mature, self-directed, and intrinsically motivated model of learning.

In summary, meaningful student learning in the era of LLMs requires a redesign of both curriculum and mindset. Educators must modernize their practices, empower students to use AI tools responsibly, and align most, if not every, aspect of teaching with the fundamental goals of education: to foster understanding, critical thinking, and lasting knowledge that students can carry with them beyond the classroom.

7.5 How can educators in higher education adapt traditional course practices to maintain intended student learning outcomes while permitting the use of Large Language Model technology by students?

The widespread availability of LLM technology requires an urgent and thoughtful response from educators and institutions worldwide [10]. Students can—and increasingly will—use tools like ChatGPT, which are rapidly becoming as commonplace as conducting a Google search [16]. Banning such technologies is difficult, if not impossible [18]. In response, many educational institutions and universities have adopted an "open but cautious" stance toward the integration of AI tools in learning environments [66][12][31].

Given that these developments have emerged within just the past two years, AI in education remains a rapidly evolving field with significant gaps in knowledge and practice [11]. For this reason, we propose both a short-term and a long-term response. The short-term response focuses on adapting current educational practices to maintain effectiveness and integrity within traditional models. The long-term response calls for a broader shift in mindset—toward reimagining education itself, grounded in new principles, learner autonomy, and thoughtful integration of emerging technologies.

Short-term adaptations

In the short term, educators can adapt traditional courses by building a robust instructional framework that supports learning in an era of information overload, misinformation, and the widespread use of generative AI. This framework can rest on four key pillars:

- **Clarify the learning goals:** Clearly define what students should understand, be able to do, or apply by the end of a course or unit. These goals should form the foundation for all instructional and assessment decisions.
- **Connect goals to future usefulness:** Help students see the relevance of what they are learning by explicitly linking course goals to real-world, academic, and professional contexts. When students understand the "why" behind their learning, they are more likely to engage meaningfully, even when AI tools are available.
- **Design assessments that align with goals:** Traditional assessments must be reevaluated and aligned directly with the intended learning outcomes. Assessments should prioritize critical thinking, application, synthesis, and original insight—skills that LLM cannot easily replicate or should also be possible without LLM use.
- **Test understanding authentically:** Shift away from relying solely on take-home essays and written assignments that can be completed with minimal effort using LLMs. Instead, use authentic assessment formats such as oral examinations, in-class tasks, presentations, practical demonstrations, or structured knowledge checks. These methods make it more difficult for students to rely passively on AI tools, allowing educators to assess their actual understanding more accurately.

Additionally, AI literacy should be integrated into the curriculum itself [82]. Students must learn how to use LLMs as tools rather than as automated, flawed answer machines [16]. This education includes instruction in topics like AI ethics and responsible use [10], Prompt engineering [93], the limitations of AI systems (e.g., hallucinations, false confidence, outdated or fabricated information), and the importance of validating AI-generated content against credible sources [28].

Learning occurs through active cognitive engagement, which involves thinking, doing, and reflecting [50]. If students rely entirely on AI to generate answers, they miss out on this essential process. Educators must, therefore, teach students how to utilize AI to support their thinking, rather than replace it. This might call for a change in how we see education today.

Moreover, current course components that are particularly vulnerable to LLM use must be critically reviewed. Some should be removed, while others should be adapted or replaced. In other cases, it remains unclear or no change is necessary. These include:

Overview of course component adaptations [27][36][35]

Category	Type	AI Perform Possible	Action	Explanation
Assessment Components	Examinations (written)	Partially	Adapt	Should be conducted offline or within a secured online environment to prevent AI-assisted responses.
""	Examinations (Oral)	No	Nothing	Remains largely unaffected by AI tools.
""	Presentations	Partially	Adapt	Allow the use of AI for preparation, but focus assessment on the student's critical understanding and ability to explain content.
""	Assignments	Yes	Adapt	Design with an AI-first approach, accepting AI use while ensuring learning outcomes are still demonstrated.
""	Projects	Yes	Adapt	Apply an AI-first mindset, recognizing that final products alone no longer prove individual skill or knowledge.
""	Portfolios	Yes	Adapt	Use as evidence of process and strategy rather than as sole proof of skill—focus on how outcomes were achieved.
""	Peer Assessment	No	Nothing	Remains a valid method, as it relies on interpersonal evaluation and judgment.
Learning Activities	Group Work	Partially	Adapt	May be influenced by AI-generated contributions; emphasize individual accountability and collaborative reflection.
""	Simulations	Partially	Unclear	Further research is needed.
""	Research	Yes	Adapt	Emphasize fact-checking, source evaluation, and transparency in process; include both research approach and critical reflection in assessment.
""	Feedback	Partially	Adapt	Teach students how to assess the reliability of AI-generated feedback; ensure essential feedback is still provided by qualified instructors.
Homework	Overall	Partially	Adapt	For all types of homework, the same principle applies: it should function as a supportive tool for learning, not merely as a task to be completed.
""	Math	Partially	Adapt	There should be a clear shift toward emphasizing its role in reinforcing understanding and encouraging students to approach it with intrinsic motivation.
""	Programming	Partially	Adapt	To support this, critical components of homework could be integrated into school hours, ensuring that students engage with the material in a structured environment with guidance.
""	Writing	Yes	Adapt	This approach also helps prevent homework from being perceived as an isolated chore and instead
""	Languages	Yes	Adapt	frames it as a meaningful part of the learning process.

Table 7.1: Overview of course components and how they can be adapted

Ultimately, if educators cannot assess the validity of a learning activity or output, it should be redesigned or replaced. Every component of a course should trace back to a clearly defined, testable learning objective. By focusing on what can truly be learned and demonstrated, educators can ensure that student learning remains meaningful, even in the presence of LLMs.

Common principles and emerging guidelines

Various organizations, institutions, and universities have begun formulating frameworks, policies, and guidelines to address the integration of AI in education. The literature reviewed reveals a growing consensus on certain core principles that appear consistently across institutional approaches. While specific policies vary depending on context, several themes have remained stable over time and are broadly reflected in most frameworks. These findings align closely with the outcomes and recommendations of this study [18][10][59][66][79][31][94][74][33][95][65].

Below is a summary of the most strongly represented and recurring policy directions observed in the literature, each accompanied by a brief explanation. These are the key principles found in institutional AI policies:

- **Human-Centered Approach:** Policies emphasize that AI tools should support, not replace, human learning and teaching. The focus remains on promoting meaningful cognitive engagement and personal development.
- **Pedagogically Grounded Decision-Making** The use of GenAI should be guided by sound didactic reasoning. Decisions to implement AI tools should be based on their contribution to learning outcomes, rather than relying solely on technological trends.
- **Decentralized Implementation for course- or Program-Level Autonomy** Many institutions encourage flexibility by allowing individual teachers, courses, or programs to set specific rules for AI use. This acknowledges that context-specific applications are often necessary for effective integration.
- **Alignment with Labor Market and Professional Practice** As AI rapidly advances, it is becoming evident that it is starting to outpace humans in critical areas such as reading, mathematics, and scientific reasoning. Educational programs are viewed as reflections of the professional fields they prepare students for. Consequently, the responsible use of AI is being incorporated into curricula to mirror evolving workplace expectations.
- **Focus on Demonstrable Learning Outcomes** Despite the availability of GenAI tools, students must still be able to demonstrate their understanding and mastery of core skills and knowledge. Assessment practices are being revised to uphold this principle.
- **Transparency in AI Use** Both students and teachers are encouraged—or required—to be transparent about their use of AI tools. This includes disclosure in assessments and classroom practices to support academic honesty and clarity.
- **Academic and Data Integrity** Safeguards are being developed to ensure that the use of AI tools does not compromise academic standards or the protection of sensitive educational data.
- **Leveraging AI to enhance education and learning** Use AI as a support tool. Many policies frame GenAI as a tool for academic support—useful for brainstorming, structuring ideas, or receiving formative feedback—but not as a substitute for original work or deep learning.
- **Educating for AI Literacy** Both students and educators must be taught how to use AI tools effectively, ethically, and critically. This includes understanding how to craft effective prompts, check for false or biased outputs, and validate the accuracy of information. AI literacy is increasingly viewed as a foundational competence for modern education.

As this paper was being finalized, the University of Amsterdam (UvA) updated its institutional AI policy, evolving from an "open but cautious" position toward a more structured framework [95]. This updated policy outlines more straightforward guidelines on how educators and students should use AI responsibly in education, reflecting many of the same principles highlighted above. These developments have been incorporated into the discussion to ensure relevance and alignment with current institutional directions.

Long-term continuation of the proposed adaptations

In the long term, adapting traditional courses to the presence of LLM technology requires a fundamental shift not only in how we teach but in how we think about the purpose and design of education itself. Rather than attempting to preserve existing methods exactly as they were, educators should recognize this moment as an opportunity for systemic renewal. The goal is not simply to replicate past practices under new constraints, but to develop more effective, future-proof approaches to teaching and learning that reflect the realities of a rapidly changing world.

Education as Preparation for a Changing Future

We no longer teach the same curriculum we did 50 years ago, and it is unlikely that today's curriculum will be relevant 50 years from now. Education must reflect the world students are preparing to enter. This reflection involves re-evaluating which subjects and skills remain essential in an era where AI can perform many routine or information-based tasks. Suppose tools like ChatGPT can generate summaries, code snippets, or even creative writing. In that case, educators must consider whether traditional instruction in those areas still holds the same value or whether the focus should shift toward using, understanding, and improving such tools.

At the same time, we must ask: what competencies must students retain even when AI is available? What if these tools are no longer accessible, due to policy, cost, or platform instability? These questions highlight the importance of striking a balance between AI integration and the development of enduring human skills, such as reasoning, creativity, ethical judgment, and effective communication.

Re-imagining Curriculum Design and Learning Goals

This shift encourages us to rethink not only what we teach, but how we teach. Effective education does not require adherence to traditional formats if newer, more relevant approaches serve the same or greater purpose. The following long-term strategies can help reshape educational practices:

- Redefine learning goals: Focus on understanding, adaptability, and the ability to apply knowledge in new contexts.
- Connect education to the real world: Ensure that what students learn aligns with future academic, personal, and professional applications.
- Design AI-resilient assessments: Prioritize performance-based, oral, and iterative evaluations that better reflect actual skill levels and discourage surface learning.
- Teach AI literacy as core content: Include topics like prompt engineering, AI limitations, AI hallucinations, and ethical use. Students should learn that asking the right questions and validating AI responses are essential to critical use.

Shifting Responsibility and Mindset

A meaningful long-term transformation also involves a shift in responsibility from the teacher to the student. Students must be encouraged to take ownership of their learning, not because they are told to, but because they see the value in it. Intrinsic motivation becomes central. In this model, assignments and homework are not mandatory by default, but optional as a learning tool. Students complete them because they want to understand the material and/or succeed on a test or assignment.

To support this shift, assessment systems must also evolve. Rather than hinging student success on high-stakes, one-time exams, we can move toward mastery-based assessment, where the default is that it is okay to retake tests or demonstrate knowledge through multiple formats until the required level is met. This approach reduces pressure, promotes learning over performance, and allows students to grow at their own pace.

Long-term adaptation to LLMs requires more than new rules—it requires a new mindset. By reassessing what we teach, updating how we assess, and encouraging students to learn for themselves rather than for grades, educators can ensure that learning remains deep, relevant, and human—even in the age of AI.

7.6 Limitations and future research

While this study offers insights into how educators might adapt traditional courses in the age of LLM technology, several limitations must be acknowledged, and important avenues for future research should be highlighted.

First, resource constraints in educational settings present a significant limitation to the practical implementation of some of the recommendations in this study. While the proposed strategies for curriculum redesign, authentic assessment, and AI literacy integration are theoretically sound, many educators face time, training, and institutional limitations that make such changes difficult to realize at scale.

Second, this study does not critically examine the effectiveness of traditional educational practices themselves. Some practices may have limited pedagogical value, regardless of the presence of LLMs. However, such a critique falls outside the scope of this work. A more thorough analysis of which teaching methods are truly effective would provide a stronger foundation for redesign efforts.

Third, there is a persistent gap between theory and practice. While the study outlines long-term strategic shifts and short-term adaptation methods, the complexity of real-world classrooms, institutional policies, and student behavior may complicate the application of these ideas. Translating theory into practice remains a major challenge that requires iterative testing and feedback.

Fourth, limitations of this study include its focus on qualitative data from a limited number of institutions and educators. The findings may not be generalizable across different national or cultural contexts, especially in regions where access to AI is limited or regulated differently.

Additionally, many of the proposals in this thesis are conceptual in nature. We based these proposals on existing literature, policy reports, and emerging practices, but empirical validation is limited. There is a pressing need for more experimentation, pilot programs, and case studies to test the effectiveness of LLM-informed teaching models, AI-integrated assessments, and student support systems in diverse educational contexts. Another limitation lies in the measurement of learning outcomes. Traditional evaluation methods may no longer capture the depth or authenticity of student understanding in AI-rich environments. Developing better tools and frameworks for assessing meaningful learning, particularly in the presence of AI assistance, should be a priority for future work.

Finally, while this study uses the Visible Learning framework as a pedagogical foundation, the research could benefit from a more grounded analysis through the lens of educational theory and empirical educational research. A deeper exploration of learning science, cognitive load theory, or self-regulated learning would enhance the robustness of the recommendations and better align them with existing educational knowledge.

This research serves as a starting point for understanding the intersection between AI technologies and traditional educational design. Future studies should:

- Conduct empirical investigations into how LLMs influence different types of learning outcomes across disciplines and educational levels.
- Explore student attitudes, motivations, and behavior when LLMs are used openly as learning tools.
- Develop and test assessment models that are resilient to AI assistance, including oral, process-based, and performance assessments.
- Investigate the feasibility and scalability of AI-integrated teaching strategies in under-resourced or high-variability learning environments.
- Extend this work through longitudinal studies, capturing the evolving role of AI in education and how student learning adapts over time.

As AI continues to transform how information is accessed, processed, and applied, education systems must evolve accordingly. This thesis contributes to that evolution by outlining practical and strategic pathways forward, while also highlighting the complexity and ongoing nature of the challenge. Continued, interdisciplinary research is essential to building a robust, equitable, and future-ready educational system in the age of LLMs.

Threats to validity

This study, while aiming to provide meaningful insights into the integration of LLM technology in traditional education settings, is subject to several potential threats to validity that should be acknowledged.

One limitation affecting internal validity is the reliance on secondary data and literature. As much of the study is based on emerging research, institutional reports, and pilot programs, the conclusions drawn are dependent on the accuracy and completeness of those external sources. Additionally, due to the rapidly evolving nature of LLM technology, findings from recent studies may quickly become outdated, which can affect the internal consistency of the conclusions. There may also be biases in interpretation, especially where qualitative sources or policy documents are involved. Researcher bias could influence how certain developments are framed, especially in assessing whether specific course components are at risk or in need of redesign.

The generalizability of the findings is another concern. Educational systems vary widely across countries, institutions, and disciplines. This study may be most applicable to higher education or digitally mature educational contexts, and its recommendations might not transfer well to under-resourced schools, early education, or non-Western educational systems. Additionally, much of the discussion assumes a baseline level of digital literacy and access to AI tools, which may not be the case for all student populations. As a result, conclusions about how LLMs can or should be integrated may not be universally applicable.

There is also a risk related to defining key constructs, such as "effective learning," "adaptation," and "AI integration." These terms are inherently broad and context-dependent. Different stakeholders, like educators, policymakers, and students, may interpret them differently. Although the study attempts to clarify these constructs through literature review and context-specific framing, some ambiguity remains.

Given that the study was conducted during a period of rapid technological change, there is a high risk to temporal validity. New capabilities, policies, or cultural shifts surrounding LLMs could rapidly alter how AI tools are utilized and perceived in education. Therefore, some findings and recommendations may be time-sensitive and require periodic re-evaluation.

Chapter 8

Conclusion

This thesis explored the integration of AI and extensive language models (LLMs) into traditional educational settings. Based primarily on qualitative analysis, this research aimed to understand how these tools are currently being used, what challenges and opportunities they present for educators and students, and how educational systems can adapt to their growing presence.

8.1 Answer to the research question

The goal was to answer the question: How can educators in higher education adapt traditional course practices to maintain intended student learning outcomes while permitting the use of Large Language Model technology by students?The study found that generative AI tools offer significant potential to enhance learning experiences. They can support personalized instruction, automate aspects of assessment, and promote student engagement through novel content creation. However, the findings also reveal substantial challenges, including issues of academic integrity, unequal access, disrupting traditional methods and a lack of clear pedagogical frameworks for effective use. Educators and students expressed both enthusiasm and concern, highlighting the need for targeted professional development and clearer institutional guidance.

From a policy perspective, this research indicates that educational systems must move beyond basic regulation and develop comprehensive strategies that meaningfully integrate generative AI into curriculum design, teacher training, and student skill development. Without such efforts, schools risk either losing potentially valuable educational tools or implementing them ineffectively. In response to the research question (How can educators in higher education adapt traditional course practices to maintain intended student learning outcomes while permitting the use of Large Language Model technology by students?), the study found that both the literature and the collected data indicate the need for a dual approach, combining short-term and long-term strategies, that involves rethinking current educational models to maintain their effectiveness in the era of AI. In the short term, a "patch" strategy is necessary: adapting existing educational practices to accommodate AI usage while preserving familiar structures and maintaining continuity. This includes adjusting assessment methods, classroom activities, and academic expectations to ensure they remain meaningful even when students have access to AI tools. In the long term, a more transformational vision emerges. This vision involves shifting toward a curriculum that mirrors the evolving demands of the workforce, places greater responsibility on students, and has its primary focus on intrinsic motivation. The long-term model emphasizes engaging, interactive learning experiences that go beyond task completion, aiming instead to make education more dynamic, relevant, and enjoyable.

8.2 Reflection

AI is a truly fascinating field, and the emergence of LLM represents one of the most impressive, impactful, and disruptive technological breakthroughs of the past decade. This shift is influencing nearly every sector, signaling a broader transition into a new digital era. Among the many affected domains, education stands out as a particularly complex and far-reaching area of impact. Understanding how AI transforms teaching, learning, and assessment is an immense topic—one that requires not only extensive academic inquiry through papers and theses, but also substantial practical data on how students perform and feel within this changing landscape.

For this research to contribute meaningfully to how education can respond to a changing landscape shaped by AI, the most effective approach was to synthesize current knowledge, identify patterns and gaps, and provide guidance for future research and experimentation. The original hypothesis (that educators could adapt traditional courses by redefining assessment practices, emphasizing higher-order cognitive skills, and redesigning instruction to incorporate productive LLM use) was in the right direction. However, it proved to be too simple for a topic that was far more complex than initially assumed. The integration of AI in education is a multidimensional challenge, where changes in one area trigger changes in others. This research offers a structured overview of available data and literature, highlights key findings, and proposes directions for future exploration and real-world testing.

8.3 Future recommendations

The field of AI in education (AIED) is closely tied to real-world practice, and as such, research in this domain should reflect that practical orientation. Theoretical assumptions alone are insufficient; statements and conclusions must be tested in real-world educational contexts. The outcomes of these implementations should serve as a foundation for new literature, ensuring that academic work is grounded in evidence rather than speculation.

It is also noteworthy that institutions across the globe are facing remarkably similar challenges related to AI and are independently pursuing similar solutions. In both the literature and interviews conducted for this research, the same concerns were repeatedly mentioned, highlighting a shared global issue. These concerns strongly suggest the need for increased collaboration between governments, educational institutions, and research bodies. Currently, numerous studies are being conducted in parallel, often without awareness of one another, resulting in fragmented progress and duplicated efforts.

Future research should explore the long-term impact of LLM use on student learning, particularly in areas such as critical thinking, creativity, and academic integrity. Large-scale quantitative studies and longitudinal research will be essential to deepen our understanding of how generative AI shapes learning outcomes over time and across diverse educational contexts.

Chapter 9

Acknowledgements

First and foremost, I would like to express my gratitude to Ana Oprescu and Adam Belloum for taking on the responsibility of co-supervising my thesis. Their guidance and advice consistently pointed me in the right direction, helping me complete my work within the academic year. I am also thankful to the programme committee, of which we are all members, for providing me with a thesis topic that I found both genuinely fascinating and intellectually rewarding.

I would like to extend my appreciation to Hans Dekkers. His lectures challenged us to think honestly, critically, and to approach life with a healthy degree of skepticism. I also want to express my admiration for his dedication, including joining our early morning workouts before his lecture and our late borrels after his lectures.

In addition, I want to take a moment to thank some amazing people simply because I can: Dr. Niko, Christina, Kees, and Robert (also referred to as Brobert or Robér) for their friendship and the many unforgettable adventures in Amsterdam; Pim and Joran for always supporting me and sticking with me since high school; Jonatan Hedin for our friendship and for inspiring me to pursue this master's; Kristina Markulin—what a life; Peter van der Meer and Michel Kruiswijk for their support, advice, and friendship; Yordi Dek for all the times we nearly died but didn't; Every member of the ACC and Erik Lopez for being both an absolute legend and an inspiring human being: to infinity.

Special thanks go to Evelyne Hillegonda Elisabeth Schoon and Bas Jacobus Kooter, not only for being an inexhaustible source of support and guidance throughout the completion of this thesis, but also during my entire master's program, and perhaps just in life overall. Your help, positive energy, and unwavering friendship makes it possible for me to persevere through everything.

I would also like to express my gratitude to Esmée for her unwavering support, the many adventures we shared, and her invaluable help during the thesis period.

I am deeply grateful to Cora Koffeman and Adriaan Thomas Zijl, who supported me in every way they could throughout my academic journey. Their encouragement and support were essential to the completion of both this thesis and my degree. A special thanks also to Eva Zijl and Naomi Zijl, who both have remarkable potential and are inspiring individuals. I look forward to seeing them put that potential to use—hopefully for more than just finding new ways to annoy me. My thanks as well to Ron Haver for all his support, advice and pressing the golden button, and to Arnoud for his welcoming attitude and encouragement.

Finally, I would like to express my heartfelt appreciation to Bertie Koffeman and Reinier Koffeman, whose door is always open and whose support is unconditional. A special thank you to Reinier Koffeman, who has always been a personal hero and someone I deeply look up to.

Chapter 10

Annotated Bibliography

This annotated bibliography serves an administrative purpose and is intended to help both the reader and the writer easily track and review the sources used in this work. It also provides insight into why specific sources were selected and the context in which they are applied. To assess the credibility of each source, a simple point-based rating system is used. Each criterion earns the source one star ★, with a maximum of ten stars. If a criterion is not met, it is marked with an empty star ☆. The rating reflects the overall quality of the source, based on the following characteristics:

★ Well-known and respected publisher
★ Well-known and respected author
★ Peer-reviewed publication
★ Modern source (2000 or later)
★ Recently published (2020 or later)
★ Significant citation count (between 50 and 200 citations)
★ High citation count (200+ citations)
★ Methodology adheres to scientific standards (based on personal interpretation)
★ Extra star if this work is outstanding
★ Paper in the same field as this paper (foundational papers)
☆ Item is missing or unknown

For example, if a well-known and respected publisher publishes a source, it is a recent publication (post-2000), and has a high citation count (over 200 citations), the rating would be as follows:

Score: ★☆☆★☆★★☆☆☆ (4)

This scoring system is designed to provide a transparent and consistent method for evaluating source quality within the context of this research. This does not conform to an official standard, but rather represents the writer's interpretation.

10.1 Literature

Title: "Why Education Matters"
Author: Kingston *et al.*
Source: [8]
Score: ★★★★☆★★★☆☆ (7)
Annotations: The authors provide insights into the outcomes of education from various perspectives, explaining why these outcomes are beneficial and important. They lay a foundation for understanding the significance of education and emphasize the importance of maintaining educational systems.
Abstract: In this article, the authors assess why educational attainment is associated with many diverse social outcomes. Their multivariate models incorporate linear (years of schooling) and non linear (credentials) measures of schooling, socioeconomic status (origin and destination), and cognitive ability. The outcome variables include attitudes toward civil liberties and gender equality, social and cultural capital, and civic knowledge. The results indicate only modest evidence of "credential effects." The mediating impacts of both cognitive ability and socioeconomic status (original and destination) are often substantial but even together do not account for all apparent "educational effects."

Title: "Importance of Education in Human Life: a Holistic Approach"
Author: Bhardwaj
Source: [6]
Score: ★☆★★☆★★☆☆☆ (4)
Annotations: The source highlights all major aspects of why education is important on both an individual and societal level. It offers perspectives from different angles and compares information on the topic using data from other countries as well. Note that while this is a useful reference, it is not a strong academic source. The references are poorly cited, and it is difficult to find information about the professor.
Abstract: The present research paper clarifies the importance of education with a holistic approach. With the advancement of science and technology the human civilization so far has travelled a long way with so many success stories of development in its hand. Whether be personal, social, political, economic or cultural development, the role of education can't be underestimated. Today we have education on every aspects of life which paves the way for the holistic development of the individual, society and the nation. Education is of course, a boon for the human life. But it is equally true that the education should not only be learning concept. Fundamentally, it should be ensuring values imparted to all students and of course, value based education is very important if we want to have good citizens. So in our entire curriculum, it should be ensured that value based education is incorporated into the chapters in the interesting way so that the children and students learn it at every step of their schooling. Value based education plays a pivotal role in the holistic development of students as the results of several such experiments have shown.

Title: "The Effects of Education as an Institution"
Author: Meyer
Source: [7]
Score: ★★★★☆★★★☆☆ (7)
Annotations: A paper from 1977 telling more about not the individual impact of students, but how education affects society from a macro-sociological perspective.
Abstract: Education is usually seen as affecting society by socializing individuals. Recently this view has been attacked with the argument that education is a system of allocation, conferring success on some and failure on others. The polemic has obscured some of the interesing implications of allocation theory for socialization theory and for research on the effects of education. But allocation theory, too, focuses on educational effects on individuals being processed. It turns out to be a special case of a more general macrosociological theory of the effects of education as a system of legitimation. Education restructures whole populations, creating and expanding elites and redefining the rights and obligations of members. The institutional effects of education as a legitimation system are explored. Comparative and experimental studies are suggested.

Title: "The Societal Consequences of Higher Education"
Author: Schofer *et al.*
Source: [96]
Score: ★★★★★★★★★☆ (9)
Annotations: What is the effect of many students in higher education? What is actually learned in higher education? This papers delves deeper into this topic. Note the citation: "Higher education was considered a consumption good, not a useful investment, and—in the eyes of some—a waste of time and resources. Even today, we wonder if much is learned in college". And "Higher education underlies key features of the contemporary world, sustaining globalization, new kinds of

societal mobilizations, and new conceptions of the economy."

Abstract: ... In short, many features of the contemporary world arise from the growing legions of people steeped in common forms of higher education. Panel regression models of contemporary cross-national longitudinal data examine these relationships. We find higher-education enrollments are associated with key dimensions of rationalization, globalization, societal mobilization, and expansion of the service economy. Central features of modern society, often seen as natural, in fact hinge on the distinctive form of higher education that has become institutionalized worldwide.

Title: "Research on Globalization and Education"
Author: Spring
Source: [48]
Score: ★★★★☆★★★★☆ (8)
Annotations: Good source on defining globalization, the definition of education and the effects of combining these major topics. It also provides more insights on institutions affecting local educational practices and policies.
Abstract: Research on globalization and education involves the study of intertwined worldwide discourses, processes, and institutions affecting local educational practices and policies. The four major theoretical perspectives concerning globalization and education are world culture, world systems, postcolonial, and culturalist. The major global educational discourses are about the knowledge economy and technology, lifelong learning, global migration or brain circulation, and neoliberalism. ...

Title: "Globalisation, Knowledge Economy and Comparative Education"
Author: Dale
Source: [97]
Score: ★★★★☆★★★★☆ (8)
Annotations: Noting to add to the abstract in this case.
Abstract: This paper seeks to introduce this special issue by setting out what seem to be some of the major theoretical and methodological issues raised for comparative education by the increasing prominence of the discourses of the knowledge economy, which, it is argued, represent a particularly strong version of globalisation and its possible relationships to education systems, and hence an especially acute challenge to comparative education. It focuses on the possible implications of these changes for each of the three elements of 'national education system'. In terms of the 'national' it discusses the nature and consequences of methodological nationalism, and emphasises the emerging pluri-scalar nature of the governance of education. In terms of 'education', it argues that education is now being asked to do different things in different ways, rather than the same things in different ways. In terms of 'system', it is suggested that the constitution of education sectors may be in the process of changing, with a development of parallel sectors at different scales with different responsibilities. Overall, the article suggests that we may be witnessing the development of a new functional, scalar and sectoral (non zero sum) division of the labour of educational governance. Finally, it addresses the question 'what is now to be compared' and considers the consequences for both 'explaining' and 'learning' through comparative education.

Title: "A Comparative Analysis of the Efficiency of National Education Systems"
Author: Thieme *et al.*
Source: [98]
Score: ★★★★☆☆☆★☆☆ (5)
Annotations: Assess the state of different education systems. Also introduces PISA (Programme for International Student Assessment) which could be an interesting source for other attributes of this research.
Abstract: The present study assesses the performance of 54 participating countries in PISA 2006. It employs efficiency indicators that relate result variables with resource variables used in the production of educational services. Desirable outputs of educational achievement and undesirable outputs of educational inequality are considered jointly as result variables. A construct that captures the quality and quantity of educational resources consumed is used as resource variables. Similarly, environmental variables of each educational system are included in the efficiency evaluation model; while these resources are not controllable by the managers of the education systems, they do affect outcomes. We find that European countries are characterized by weak management, the Americans (mainly Latin Americans) by a weak endowment of resources, and the Asians by a high level of heterogeneity. In particular, Asia combines countries with optimal systems (South Korea and Macao-China); countries with managerial problems (Hong Kong, China-Taipei, Japan and Israel); others where the main challenge is the weak endowment of resources (Jordan and Kyrgyzstan), and, finally, others where the main problem is in the long run since it concerns structural conditions of a socioeconomic and cultural nature

(Turkey, Thailand, and Indonesia).

Title: "Education Systems of Countries in Comparison"
Author: Qalandarova and Iskandarova
Source: [99]
Score: ★☆☆★☆☆☆☆☆☆ (2)
Annotations: Article that show what I want to find in literature, the article itself is not a strong academic source. Don't reference this one, or explain why you reference it when you do.
Abstract: This article compares education systems in different countries around the world. This work analyzes the educational processes of developed and developing countries, the factors affecting the quality of education, and the main differences between these systems. The possibility of learning from international experience is considered by studying the differences in education systems between countries.

Title: "How Does Philippines's Education System Compare to Finland's?"
Author: Kilag *et al.*
Source: [100]
Score: ★☆☆★★☆☆★☆☆ (4)
Annotations: An interesting source that compares the education systems of Finland and the Philippines, supported by strong data. It can be used for comparing different education systems worldwide.
Abstract: This study examines the differences between the education systems in Finland and the Philippines. The education systems in both countries have significant differences in terms of policies, teacher qualifications, and curriculum structure. Finland has a highly decentralized curriculum that allows teachers to develop their own curriculum according to the needs and interests of students, while the curriculum in the Philippines is centralized and designed by the central government. The quality and competence of teachers in Finland are also higher compared to those in the Philippines, where a Bachelor's or Diploma Four degree is the minimum requirement for teaching. Finland's success in education can be attributed to the government's persistence and consistency in implementing education policies formulated over the years, and their emphasis on teacher quality and competence. On the other hand, the Philippines has struggled to improve the quality of education, with some policies having little to no positive impact. Given the differences between the two education systems, the study suggests that the Philippines should learn from Finland's education system and adopt some of its policies to improve its own education system. By doing so, the Philippines can improve the quality of education and potentially achieve the same success that Finland has achieved.

Title: "Comparative Analysis of the Differences in the Education System in China vs Germany"
Author: Huang
Source: [49]
Score: ★☆☆★★☆☆★☆☆ (4)
Annotations: An interesting source that compares the education systems of China and the Germany, supported by strong data. It can be used for comparing different education systems worldwide. Similar as the previous source.
Abstract: This paper aims to point out the insufficiency of in-service teacher education in China and Germany, and correspondingly provide suggestions for enhancing teacher's quality and quantity effectively to fit the higher requirements of school education. In this study, in-service teacher education in China and Germany are discussed through a qualitative analysis. This paper resolves the following issues through the comparison of in-service teacher education between China and Germany: (1) Exploration of similarities and differences of the institutions, aims, and modes of in-service education of the two countries. (2) Analysis of the educational management measures between the two countries; finding the advantages of these measures to improve the teachers' quality and quantity. These measures include systems of admission, appointment, treatment, and assessment. (3) Point out the problems of teacher education in China and Germany, and correspondingly give some suggestions to improve on inadequacies. This paper fulfills an identified need to improve new teachers' professional development through transforming previous studies with teaching experience in real educational situations.

Title: "Role of AI chatbots in education: systematic literature review"
Author: Labadze *et al.*
Source: [2]
Score: ★★★★★★★★★★ (10)
Annotations: A paper in the same era as this one, providing a literature review of A.I. Chatbots in education. Very useful since it provides a lot of insights how these chatbots are used and what the impact is, can serve as a foundational paper for this paper.
Abstract: AI chatbots shook the world not long ago with their potential to revolutionize education

systems in a myriad of ways. AI chatbots can provide immediate support by answering questions, offering explanations, and providing additional resources. Chatbots can also act as virtual teaching assistants, supporting educators through various means. In this paper, we try to understand the full benefits of AI chatbots in education, their opportunities, challenges, potential limitations, concerns, and prospects of using AI chatbots in educational settings. We conducted an extensive search across various academic databases, and after applying specific predefined criteria, we selected a final set of 67 relevant studies for review. The research findings emphasize the numerous benefits of integrating AI chatbots in education, as seen from both students' and educators' perspectives. We found that students primarily gain from AI-powered chatbots in three key areas: homework and study assistance, a personalized learning experience, and the development of various skills. For educators, the main advantages are the time-saving assistance and improved pedagogy. However, our research also emphasizes significant challenges and critical factors that educators need to handle diligently. These include concerns related to AI applications such as reliability, accuracy, and ethical considerations.

Title: "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature"
Author: Lo
Source: [3]
Score: ★★★★★★★★★★ (10)
Annotations: Another paper that can play a key part in this paper, as it describes the impact of a large LLM systems (ChatGPT) on education. The wants to address how to change education because of LLM technology, while Lo's paper address the impact. The impact can be used to describe what needs to be changed.
Abstract: An artificial intelligence-based chatbot, ChatGPT, was launched in November 2022 and is capable of generating cohesive and informative human-like responses to user input. This rapid review of the literature aims to enrich our understanding of ChatGPT's capabilities across subject domains, how it can be used in education, and potential issues raised by researchers during the first three months of its release (i.e., December 2022 to February 2023). A search of the relevant databases and Google Scholar yielded 50 articles for content analysis (i.e., open coding, axial coding, and selective coding). The findings of this review suggest that ChatGPT's performance varied across subject domains, ranging from outstanding (e.g., economics) and satisfactory (e.g., programming) to unsatisfactory (e.g., mathematics). Although ChatGPT has the potential to serve as an assistant for instructors (e.g., to generate course materials and provide suggestions) and a virtual tutor for students (e.g., to answer questions and facilitate collaboration), there were challenges associated with its use (e.g., generating incorrect or fake information and bypassing plagiarism detectors). Immediate action should be taken to update the assessment methods and institutional policies in schools and universities. Instructor training and student education are also essential to respond to the impact of ChatGPT on the educational environment.

Title: "Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines"
Author: Wang *et al.*
Source: [4]
Score: ★☆★★★☆☆★★★ (7)
Annotations: Another paper that can play a key part in this paper, as it describes the impact of a large LLM systems (ChatGPT) on education. This wants to address how to change education because of LLM technology, while Lo's paper address the impact of LLM on education. The impact can be used to describe what needs to be changed.
Abstract: The advancements in Generative Artificial Intelligence (GenAI) can provide opportunities for enriching educational experiences, but at the same time raise concerns regarding academic integrity. Many educators have expressed anxiety and hesitation when it comes to integrating GenAI in their teaching practices. Thus, recommendations and guidance from institutions are needed to support instructors in this new and emerging GenAI era. In response to this need, this study explores different U.S. universities' academic policies and guidelines regarding the use of GenAI tools (e.g., ChatGPT) for teaching and learning, and from there, gains understanding of how these universities respond and adapt to the development of GenAI in their academic contexts. Data sources include academic policies, statements, guidelines, and relevant resources provided by the top 100 universities in the U.S. Results show that the majority of these universities adopt an open but cautious approach towards GenAI. Primary concerns lie in ethical usage, accuracy, and data privacy. Most universities actively respond and provide diverse types of resources, such as syllabus templates, workshops, shared articles, and one-on-one consultations; focusing on a range of topics, namely general technical introduction, ethical concerns, pedagogical applications, preventive

strategies, data privacy, limitations, and detective tools. The findings provide four practical pedagogical implications for educators when considering GenAI in teaching practices: 1) accepting GenAI presence, 2) aligning GenAI use with learning objectives, 3) evolving curriculum to prevent misuse of GenAI, and 4) adopting multifaceted evaluation strategies. For recommendations toward policy making, the article suggests two possible directions for the use of GenAI tools: 1) establishing discipline-specific policies and guidelines, and 2) managing students' sensitive information in a transparent and careful manner.

Title: "A systematic literature review of empirical research on ChatGPT in education"
Author: Albadarin *et al.*
Source: [11]
Score: ★★★★★☆★★★ (9)
Annotations: Albarain's paper reflects exactly why this paper is needed: it shows the up and downsides of LLM use by students, and also emphasizes that students use less in some scenarios: "... that overuse of ChatGPT may negatively impact innovative capacities and collaborative learning competencies among learners ...."
Abstract: Over the last four decades, studies have investigated the incorporation of Artificial Intelligence (AI) into education. A recent prominent AI-powered technology that has impacted the education sector is ChatGPT. This article provides a systematic review of 14 empirical studies incorporating ChatGPT into various educational settings, published in 2022 and before the 10th of April 2023—the date of conducting the search process. It carefully followed the essential steps outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines, as well as Okoli's (Okoli in Commun Assoc Inf Syst, 2015) steps for conducting a rigorous and transparent systematic review. In this review, we aimed to explore how students and teachers have utilized ChatGPT in various educational settings, as well as the primary findings of those studies. By employing Creswell's (Creswell in Educational research: planning, conducting, and evaluating quantitative and qualitative research [Ebook], Pearson Education, London, 2015) coding techniques for data extraction and interpretation, we sought to gain insight into their initial attempts at ChatGPT incorporation into education. This approach also enabled us to extract insights and considerations that can facilitate its effective and responsible use in future educational contexts. The results of this review show that learners have utilized ChatGPT as a virtual intelligent assistant, where it offered instant feedback, on-demand answers, and explanations of complex topics. Additionally, learners have used it to enhance their writing and language skills by generating ideas, composing essays, summarizing, translating, paraphrasing texts, or checking grammar. Moreover, learners turned to it as an aiding tool to facilitate their directed and personalized learning by assisting in understanding concepts and homework, providing structured learning plans, and clarifying assignments and tasks. However, the results of specific studies (n=3, 21.4 procent) show that overuse of ChatGPT may negatively impact innovative capacities and collaborative learning competencies among learners ...

Title: *Education at a Glance 2024: OECD Indicators*
Author: OECD
Source: [10]
Score: ★★★★★★★★★★ (10)
Annotations: Incredible rich data source of the OECD, yearly publish data and information on the global state of education.
Abstract: Education at a Glance is the definitive guide to the state of education around the world. More than 100 charts and tables in the publication and country notes – as well as many more in the data explorer – describe the output of educational institutions; the impact of learning across countries; access, participation and progression in education; the financial resources invested in education; and teachers, the learning environment and the organisation of schools. The 2024 edition focuses on equity in education, providing indicators on gaps in educational outcomes and discussing the effect of educational attainment on labour market outcomes.

Title: "Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education"
Author: Qadir
Source: [28]
Score: ★★★★★★★★★★ (10)
Annotations: Useful source on A.I. in education in the engineering field.
Abstract: Engineering education is constantly evolving to keep up with the latest technological developments and meet the changing needs of the engineering industry. One promising development in this field is the use of generative artificial intelligence technology, such as the ChatGPT conversational agent. ChatGPT has the potential to offer personalized and effective learning experiences by

providing students with customized feedback and explanations, as well as creating realistic virtual simulations for hands-on learning. However, it is important to also consider the limitations of this technology. ChatGPT and other generative AI systems are only as good as their training data and may perpetuate biases or even generate and spread misinformation. Additionally, the use of generative AI in education raises ethical concerns such as the potential for unethical or dishonest use by students and the potential unemployment of humans who are made redundant by technology. While the current state of generative AI technology represented by ChatGPT is impressive but flawed, it is only a preview of what is to come. It is important for engineering educators to understand the implications of this technology and study how to adapt the engineering education ecosystem to ensure that the next generation of engineers can take advantage of the benefits offered by generative AI while minimizing any negative consequences.

Title: *AI and education: guidance for policy-makers*
Author: UNESCO *et al.*
Source: [18]
Score: ★★☆★★★☆★★★ (8)
Annotations: Strong source, mainly for background information on A.I. in education.
Abstract: Within just the last five years, because of some prominent successes and its disruptive potential, artificial intelligence (Al) has moved from the backwaters of academic research to the forefront of public discussions, including those at the level of the United Nations. In many countries, Al has become pervasive in daily life - from smartphone personal assistants to customer support chatbots, from recommending entertainment to predicting crime, and from facial recognition to medical diagnoses.

Title: "The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach"
Author: Mai *et al.*
Source: [29]
Score: ★★☆★★★☆★★★ (8)
Annotations: Strong source, mainly for background information on A.I. in education.
Abstract: The integration of ChatGPT, an advanced AI-powered chatbot, into educational settings, has caused mixed reactions among educators. Therefore, we conducted a systematic review to explore the strengths and weaknesses of using ChatGPT and discuss the opportunities and threats of using ChatGPT in teaching and learning.

Title: "We Can Rely on ChatGPT as an Educational Tutor: A Cross-Sectional Study of its Performance, Accuracy, and Limitations in University Admission Tests"
Author: Beltozar-Clemente *et al.*
Source: [30]
Score: ★★★★★★☆★★★ (9)
Annotations: Field experiment of the use of chatGPT as a tutor and it reliability and performance. This study has evaluated LLM performance on academic tasks to measure its tutoring potential. This study tested ChatGPT on university entrance exam questions to see if it could serve as a competent tutor
Abstract: The aim of this research was to evaluate the performance of ChatGPT in answering multiple-choice questions without images in the entrance exams to the National University of Engineering (UNI) and the Universidad Nacional Mayor de San Marcos (UNMSM) over the past five years. In this prospective exploratory study, a total of 1182 questions were gathered from the UNMSM exams and 559 questions from the UNI exams, encompassing a wide range of topics including academic aptitude, reading comprehension, humanities, and scientific knowledge. The results indicate a significant ($p < 0.001$) and higher proportion of correct answers for UNMSM, with 72 procent (853/1182) of questions answered correctly. In contrast, there is no significant difference ($p = 0.168$) in the proportion of correct and incorrect answers for UNI, with 52 procent (317/552) of questions answered correctly. Similarly, in the World History course ($p = 0.037$), ChatGPT achieved its highest performance at a general level, with an accuracy of 91 procent. However, this was not the case in the language course ($p = 0.172$), where it achieved the lowest score of 55 procent. In conclusion, to fully harness the potential of ChatGPT in the educational setting, continuous evaluation of its performance, ongoing feedback to enhance its accuracy and minimize biases, and tailored adaptations for its use in educational settings are essential.

Title: *The Art of Computer Programming*
Author: Knuth
Source: [90]
Score: ★★☆☆☆★☆★★☆ (5)
Annotations: Donald Knuth is a pioneer in computer science and wrote many books on programming. Although he created outstanding work and

his literature can be used in many ways for this thesis, for now it is just used to help designing the background section.

Abstract: The Art of Computer Programming (TAOCP) is a detailed, multi-volume work by computer scientist Donald Knuth that explores programming algorithms and their analysis. As of 2025, volumes 1, 2, 3, 4A, and 4B have been published, with additional volumes anticipated. Volumes 1 through 5 are designed to form the foundational core of programming for sequential machines, while Volumes 6 and 7 will address more specialized topics.Knuth originally envisioned the project in 1962 as a single book with twelve chapters. The first three volumes—initially part of a planned seven-volume series—were published in 1968, 1969, and 1973. Although work on Volume 4 began in 1973, it was paused in 1977 to focus on typesetting issues related to the second edition of Volume 2. Knuth resumed writing Volume 4A by hand in 2001, and that same year saw the online release of its first pre-fascicle, 2A. The first paperback installment, Fascicle 2, came out in 2005, and the hardcover Volume 4A, compiling Fascicles 0 through 4, was released in 2011. Subsequent fascicles—Fascicle 6 on satisfiability (2015) and Fascicle 5 on mathematical preliminaries, backtracking, and dancing links (2019)—were added later.Volume 4B builds upon the content from Fascicles 5 and 6. The manuscript was submitted for publication on August 1, 2022, and the volume was published the following month. Fascicle 7, covering constraint satisfaction and intended for the upcoming Volume 4C, was the topic of Knuth's August 3, 2022 talk and was officially published on February 5, 2025.

Title: "The Anatomy of a Large-Scale Hypertextual Web Search Engine"
Author: Brin and Page
Source: [85]
Score: ★★☆☆☆★☆★★☆ (5)
Annotations: Shows the impact and workings of google during the time it gained popularity, used for the introduction / background.
Abstract: In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems.

To engineer a search engine is a challenging task. Search engines index tens to hundreds of millions of Web pages involving a comparable number of distinct terms. They answer tens of millions of queries every day. Despite the importance of large-scale search engines on the Web, very little academic research has been done on them. Furthermore, due to rapid advance in technology and Web proliferation, creating a Web search engine today is very different from three years ago. This paper provides an in-depth description of our large-scale Web search engine — the first such detailed public description we know of to date.

Apart from the problems of scaling traditional search techniques to data of this magnitude, there are new technical challenges involved with using the additional information present in hypertext to produce better search results. This paper addresses this question of how to build a practical large-scale system which can exploit the additional information present in hypertext. Also we look at the problem of how to effectively deal with uncontrolled hypertext collections where anyone can publish anything they want.

Title: "Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda"
Author: Zuiderwijk *et al.*
Source: [86]
Score: ★★★★☆★★★★☆ (8)
Annotations: About AI use and public governance. Used for the introduction.
Abstract: To lay the foundation for the special issue that this research article introduces, we present 1) a systematic review of existing literature on the implications of the use of Artificial Intelligence (AI) in public governance and 2) develop a research agenda. First, an assessment based on 26 articles on this topic reveals much exploratory, conceptual, qualitative, and practice-driven research in studies reflecting the increasing complexities of using AI in government – and the resulting implications, opportunities, and risks thereof for public governance. Second, based on both the literature review and the analysis of articles included in this special issue, we propose a research agenda comprising eight process-related recommendations and seven content-related recommenda- tions. Process-wise, future research on the implications of the use of AI for public governance should move to- wards more public sector-focused, empirical, multidisciplinary, and explanatory research while focusing more on specific forms of AI rather than AI in general. Content-wise, our research agenda calls for the development of solid, multidisciplinary, theoretical foundations for the use of AI for public governance, as well as investigations of effective implementation, engagement, and communication plans for government strategies on AI use in the

public sector. Finally, the research agenda calls for research into managing the risks of AI use in the public sector, governance modes possible for AI use in the public sector, performance and impact measurement of AI use in government, and impact evaluation of scaling-up AI usage in the public sector.

Title: *Four Arguments for the Elimination of Television*
Author: Mander
Source: [19]
Score: ☆☆★★☆★★★★☆ (6)
Annotations: Effect of the availability of TV to the general public, used for the introduction.
Abstract: Twenty-five billion dollars a year is spent in advertising, which is more than we spend on higher education nationally. The first problem we have in the information age is too much information inundating people with conflicting versions of increasingly complex events. People are giving up on understanding: overload leads to passiveness, not involvement. The second problem is that TV causes us to exchange active, first-hand experience for passive second-hand experience. That is, our experience is mediated by someone else's perception of reality.

Title: *Writing and Using Learning Outcomes: A Practical Guide*
Author: Kennedy
Source: [35]
Score: ★★★★☆★★★★☆ (8)
Annotations: Respected source that is quotes over a 1000 times about educational practices and guidelines, useful for defining education concepts and historic information.
Abstract: Given that one of the main features of the Bologna process is the need to improve the traditional ways of describing qualifications and qualification structures, all modules and programmes in third level institutions throughout the European Higher Education Area should be (re)written in terms of learning outcomes. Learning outcomes are used to express what learners are expected to achieve and how they are expected to demonstrate that achievement. This article presents a summary of developments in curriculum design in higher education in recent decades and, drawing on recent practical experience, suggests a user-friendly methodology for writing modules, courses and programmes in terms of learning outcomes.

Title: "Problem-Based Learning: What and How Do Students Learn?"
Author: Hmelo-Silver
Source: [42]
Score: ★★★☆★★★★★★ (9)
Annotations: Paper on PBL and their effectiveness, show a specific teaching approach. Strong source on this approach with many citations, interesting to dive in deeper.
Abstract: Problem-based approaches to learning have a long history of advocating experience-based education. Psychological research and theory suggests that by having students learn through the experience of solving problems, they can learn both content and thinking strategies. Problem-based learning (PBL) is an instructional method in which students learn through facilitated problem solving. In PBL, student learning centers on a complex problem that does not have a single correct answer. Students work in collaborative groups to identify what they need to learn in order to solve a problem. They engage in self-directed learning (SDL) and then apply their new knowledge to the problem and reflect on what they learned and the effectiveness of the strategies employed. The teacher acts to facilitate the learning process rather than to provide knowledge. The goals of PBL include helping students develop 1) flexible knowledge, 2) effective problem-solving skills, 3) SDL skills, 4) effec- tive collaboration skills, and 5) intrinsic motivation. This article discusses the nature of learning in PBL and examines the empirical evidence supporting it. There is considerable research on the first 3 goals of PBL but little on the last 2. Moreover, minimal research has been conducted outside medical and gifted education. Understanding how these goals are achieved with less skilled learners is an important part of a research agenda for PBL. The evidence suggests that PBL is an instructional approach that offers the potential to help students develop flexible understanding and lifelong learning skills.

Title: "Inductive Teaching and Learning Methods: Definitions, Comparisons, and Research Bases"
Author: Prince and Felder
Source: [23]
Score: ★★★☆★★★★★★ (?)
Annotations:
Abstract:

Title: "Measuring Learning Outcomes in Higher Education"
Author: Liu

Source: [51]
Score: ★★☆★★☆☆★★☆ (6)
Annotations: Highlights the importance of accountability and quality assurance in higher education, coming from a well respected source.
Abstract: How do we know what students have learned after they have been in college for four years or even longer? As college intuitions and fees continue to grow, students, parents and public policymakers are interested in understanding how public universities operate and whether their investments are well utilized. Accountability in public higher education has come into focus following the attention accountability has received in K-12 education.

Title: "Educational effectiveness research (EER): a state-of-the-art review"
Author: Reynolds *et al.*
Source: [52]
Score: ★★★★☆★☆★☆☆ (6)
Annotations: Research on EER, not research but still a solid overview on educational practices and effectiveness.
Abstract: Research and scholarship into educational effectiveness research (EER) is comprehensively reviewed from the UK, The Netherlands, the US, Cyprus, Belgium, Sweden,France, Germany, New Zealand, Australia, and other societies, dating from the field'sorigins in the 1970s. Issues include its history, methodological and theoreticaladvances, scientific properties of school effects, processes at school and classroomlevel behind these effects, the somewhat limited translation of findings into policy andpractice across the world, and future directions for research and practice in EER andfor all of the discipline more generally. Future research needs are argued to be a furtherconcentration upon teaching/teachers, more longitudinal studies, more work on possi-ble context specificity, exploration of the cross-level transactions between schools andtheir teachers/-classrooms, the adoption of "efficiency" as well as "effectiveness" asoutcome measures, and a renewed focus upon the education of the disadvantaged, theoriginal focus of our discipline when it began.

Title: *Methods for Effective Teaching: Meeting the Needs of All Students*
Author: Burden and Byrd
Source: [88]
Score: ★★☆★☆★★☆★☆ (6)
Annotations: Extensive literature on effective teaching, covering all topic related to this paper. Can serve as a foundation for key topics in education.
Abstract: he eighth edition of Methods for Effective Teaching provides research-based coverage of general teaching methods while emphasizing contemporary topics such as culturally responsive teaching, differentiated instruction, and data-driven decision making. The numerous features, tables, and lists of recommendations ensure that the text is reader friendly and practically oriented. Its unique content includes strategies to promote student understanding, differentiate instruction, manage lesson delivery, apply motivational techniques for instruction and assessment, and work with colleagues and parents. In addition, thorough coverage of classroom management and discipline is provided, along with ways to create a positive learning environment.

Title: *Visible Learning: The Sequel: A Synthesis of Over 2,100 Meta-Analyses Relating to Achievement*
Author: Hattie
Source: [24] [22]
Score: ★★★★★★★★★★ (10)
Annotations: Extensive literature on effective teaching and student learning, combining many data and paper in one book. It can serve as a foundational overview of effective education, supported by other papers.
Abstract: This unique and ground-breaking book is the result of 25 years research and synthesises over 800 meta-analyses on the influences on achievement in school-aged students. It builds a story about the power of teachers, feedback, and a model of learning and understanding. The research involves many millions of students and represents the largest ever evidence based research into what actually works in schools to improve learning. Areas covered include the influence of the student, home, school, curricula, teacher, and teaching strategies. A model of teaching and learning is developed based on the notion of visible teaching and visible learning.

Title: *How People Learn: Brain, Mind, Experience, and School*
Author: Bransford *et al.*
Source: [24] [38]
Score: ★★★★★★★★★★ (10)
Annotations: Highly impactful and cited literature on how people learn, based on a two-three year study of two committees of the Commission on Behavioral and Social Sciences and Education of the National Research Council (NRC) back in the early 2000's.

Abstract: This expanded edition of How People Learn is the result of the work of two committees of the Commission on Behavioral and Social Sciences and Education of the National Research Council (NRC). The original volume, published in April 1999, was the product of a 2-year study conducted by the Committee on Developments in the Science of Learning. Following its publication, a second NRC committee, the Committee on Learning Research and Educational Practice, was formed to carry that volume an essential step further by exploring the critical issue of how better to link the findings of research on the science of learning to actual practice in the classroom. The results of that effort were captured in How People Learn: Bridging Research and Practice, published in June 1999. The present volume draws on that report to expand on the findings, conclusions, and research agenda presented in the original volume. During the course of these efforts, a key contributor and one of the most eloquent voices on the importance of applying the science of learning to classroom practice was lost. The educational community mourns the death of Ann L. Brown, Graduate School of Education, University of California at Berkeley, cochair of the Committee on Developments in the Science of Learning and an editor of How People Learn. Her insight and dedication to improving education through science will be sorely missed.

Title: "'The playing-exploring child': Reconceptualizing the relationship between play and learning in early childhood education"
Author: Nilsson *et al.*
Source: [40]
Score: ★★★★☆★☆★☆☆ (6)
Annotations: Addressing play in early childhood and the impact on the view of education on this topic.
Abstract: In this article, the authors problematize the dichotomization of play and learning that often shapes the agenda of early childhood education research and practice. This dichotomization is driven in part by the tendency to define learning in terms of formal learning (i.e. learning as an outcome of direct instruction and school-based approaches that focus on teacher-led, goal-directed activities and declarative knowledge). The authors argue for a reconceptualization of early childhood education that understands learning and development not as an outcome, primarily, of instruction and teaching, but as an outcome of play and exploration. They develop this argument by drawing on Vygotsky's theories of play, imagination, realistic thinking and creativity. These theories challenge another dichotomy – that between imagination and reality – by arguing that imagination is implicated in the meaning-making of both play and exploration. Instead of relating play to learning where play is characterized by imagination and learning by reality, the authors' reconceptualization relates play to exploration and proposes that learning, defined as leading to human development, is an outcome of both of these activities. The authors further develop their argument by presenting ethnographic material from a qualitative research project implemented in three Swedish preschools whose practices are influenced by the Reggio Emilia pedagogical approach. The research conducted in this study contributes to new perspectives on the relationship between play and learning by introducing exploration as a counterpart to play,

Title: "Learning Through Play at School – A Framework for Policy and Practice"
Author: Parker *et al.*
Source: [39]
Score: ★★★★★★★★★★ (10)
Annotations: Journal of Teacher Education, backed by LEGO, and reviewed by several experts. Good scientific method used to research the effective learning implemeting play elements, focused on how to implement this in education.
Abstract: Learning through play has emerged as an important strategy to promote student engagement, inclusion, and holistic skills development beyond the preschool years. Policy makers, researchers and educators have promoted the notion that learning though play is developmentally appropriate—as it leverages school-age children's innate curiosity while easing the often difficult transition from preschool to school. However, there is a dearth of evidence and practical guidance on how learning through play can be employed effectively in the formal school context, and the conditions that support success. This paper addresses the disconnect between policy, research and practice by presenting a range of empirical studies across a number of well-known pedagogies. These studies describe how children can foster cognitive, social, emotional, creative and physical skills through active engagement in learning that is experienced as joyful, meaningful, socially interactive, actively engaging and iterative. The authors propose an expanded definition for learning through play at school based on the science of learning, and summarize key findings from international studies on the impact of children's learning through play. They identify four key challenges that underpin the considerable gap between education policy and

practice, and propose a useful framework that addresses these challenges via a common language and structure to implement learning through play.

Title: "Learning through play: a review of the evidence"
Author: Zosh *et al.*
Source: [41]
Score: ★★★★☆★☆★★★☆ (7)
Annotations: Interesting paper backed by the Lego foundation, on implementing play elements to improve learning. More white paper like, still useful.
Abstract: The aim of the LEGO Foundation is to build a future where learning through play empowers children to become creative, engaged, lifelong learners. This ambition is more critical than ever. The world of today and tomorrow is one of challenges but also of tremendous opportunity. An increasingly interconnected and dynamic reality means children will face continuous re-skilling and a need for lifelong learning as they grow. Many children also face hardship in the shape of stress, poverty and conflict. They need positive experiences and coping skills to counterbalance negative factors in their lives, and support their confidence and opportunity for making a difference. We firmly believe that promoting children's drive to learn, their ability to imagine alternatives, and to connect with their surroundings in positive ways, is absolutely essential.

Title: "Teaching for Quality Learning at University - By John Biggs & Catherine Tang"
Author: Cowan
Source: [101]
Score: ★★★★☆★★★★★☆ (8)
Annotations: English for Academic Purposes, from the teaching perspective. Relate to the book Teaching for Quality Learning at University.
Abstract: The aim of EAP essentials is people involved in "Englishfor academic purposes" programmes—such as teach-ers, teacher educators, practitioners and researcherswho may be interested in gaining a grasp of the theoryand practice behind this multi-faceted educationalarea. It aims to cover all areas of English for academicpurposes and gives a comprehensive analysis of thekey issues in the field, with a very practical approach.In fact, it translates principles into practice by involv-ing the readers in the analysis of authentic casestudies. It includes a series of "awareness-raisingtasks" to help teachers reflect on their practices. Anaccompanying compact disc supplies useful class-room materials for eap teachers' use.

Title: *Teaching for Quality Learning at University*
Author: Biggs *et al.*
Source: [50]
Score: ★★★★★★★★★★ (10)
Annotations: Foundational piece of literature to strengthen the claims made about education and the visible learning theory, about how to provide quality education at universities.
Abstract: Teaching for Quality Learning at University by John Biggs and Catherine Tang is a foundational text in the field of higher education pedagogy. The book provides a comprehensive framework for designing and delivering effective university teaching grounded in the principles of constructive alignment. This approach ensures that intended learning outcomes, teaching methods, and assessment tasks are coherently aligned to promote deep learning. The authors draw on research-based insights and practical examples to guide educators in fostering student-centered learning environments. Key themes include curriculum design, assessment strategies, quality assurance, and the role of reflective practice. Widely used by university teachers, academic developers, and policy makers, the book serves as both a theoretical guide and a practical resource for enhancing teaching and learning in higher education settings.

Title: *The Shallows: How the Internet Is Changing the Way We Think, Read and Remember*
Author: Carr
Source: [20]
Score: ★★☆★☆★☆★☆ (5)
Annotations: View on the impact of internet, used for the background and introduction.
Abstract: Welcome to The Shallows. When I wrote this book ten years ago, the prevailing view of the Internet was sunny, often ecstatically so. We reveled in the seemingly infinite bounties of the online world. We admired the wizards of Silicon Valley and trusted them to act in our best interest. We took it on faith that computer hardware and software would make our lives better, our minds sharper. In a 2010 Pew Research survey of some 400 prominent thinkers, more than eighty percent agreed that "by 2020, people's use of the Internet [will have] enhanced human intelligence; as people are allowed unprecedented access to more information, they become smarter and make better choices.

Title: "Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for

Education”
Author: Qadir
Source: [17]
Score: ★★★★★☆★★☆☆ (7)
Annotations: About ChatGPT in engineering education
Abstract: Engineering education is constantly evolving to keep up with the latest technological developments and meet the changing needs of the engineering industry. One promising development in this field is the use of generative artificial intelligence technology, such as the ChatGPT conversational agent. ChatGPT has the potential to offer personalized and effective learning experiences by providing students with customized feedback and explanations, as well as creating realistic virtual simulations for hands-on learning. However, it is important to also consider the limitations of this technology.

Title: *How People Learn: Brain, Mind, Experience, and School*
Author: Bransford *et al.*
Source: [38]
Score: ★★★★☆★☆★☆☆ (6)
Annotations: Book about learning from the 2000’s
Abstract: This expanded edition of How People Learn is the result of the work of two committees of the Commission on Behavioral and Social Sciences and Education of the National Research Council (NRC). The original volume, published in April 1999, was the product of a 2-year study conducted by the Committee on Developments in the Science of Learning.

Title: “AI technologies for education: Recent research & future directions”
Author: Zhang and Aslan
Source: [102]
Score: ★★★★★★★★☆☆ (8)
Annotations: Current state of AI for education in 2021, before the ChatGPT rise.
Abstract: From unique educational perspectives, this article reports a comprehensive review of selected empirical studies on artificial intelligence in education (AIEd) published in 1993–2020, as collected in the Web of Sciences database and selected AIEd-specialized journals. A total of 40 empirical studies met all selection criteria, and were fully reviewed using multiple methods, including selected bibliometrics, content analysis and categorical meta-trends analysis. This article reports the current state of AIEd research, highlights selected AIEd technologies and applications, reviews their proven and potential benefits for education, bridges the gaps between AI technological innovations and their educational applications, and generates practical examples and inspirations for both technological experts that create AIEd technologies and educators who spearhead AI innovations in education. It also provides rich discussions on practical implications and future research directions from multiple perspectives. The advancement of AIEd calls for critical initiatives to address AI ethics and privacy concerns, and requires interdisciplinary and transdisciplinary collaborations in large-scaled, longitudinal research and development efforts.

Title: “State of the art and practice in AI in education”
Author: Holmes and Tuomi
Source: [45]
Score: ★★★★★★★★★☆ (9)
Annotations: Provides a solid view of Ai in education before the chatGPT overtake, but more at the start. Interesting for the background, introduction and to compare different time periods relating to AI in education.
Abstract: Recent developments in Artificial Intelligence (AI) havegenerated great expectations for the future impact of AIin education and learning (AIED). Often these expectationshave been based on misunderstanding current technicalpossibilities, lack of knowledge about state- of-the-art AI ineducation, and exceedingly narrow views on the functionsof education in society. In this article, we provide a reviewof existing AI systems in education and their pedagogic andeducational assumptions. We develop a typology of AIEDsystems and describe different ways of using AI in educa-tion and learning, show how these are grounded in differ-ent interpretations of what AI and education is or couldbe, and discuss some potential roadblocks on the AIEDhighway.

Title: *Using Evidence of Student Learning to Improve Higher Education*
Author: Kuh *et al.*
Source: [47]
Score: ★★☆★★☆★★☆☆ (6)
Annotations: Book on improving higher education
Abstract: understanding what students know and are able to do as a result of their college education is no simple task, yet it is fundamental to student success and to the quality and effectiveness of American higher education. This vol-

ume grows out of a deep concern that the practical value of otherwise well-conceived efforts to assess student learning in American higher education is often diminished by deeply nested misconceptions. Many in the academy—especially those most directly responsible for the assessment of student learning—still view the assessment of student learning as an obligatory, externally imposed chore of compliance and accountability. Yes, to be fair, the capacity and commitment of colleges and universities to assess student learning outcomes have grown substantially, especially over the last decade. But the fruits of these investments— the tangible benefit ts to students and academic institutions —are embarrassingly modest.

Title: "A SWOT analysis of ChatGPT: Implications for educational practice and research"
Author: Farrokhnia *et al.*
Source: [89]
Score: ★★★★★★★★★☆ (9)
Annotations: Impact of ChatGPT on educational practices, super useful.
Abstract: ChatGPT is an AI tool that has sparked debates about its potential implications for education. We used the SWOT analysis framework to outline ChatGPT's strengths and weaknesses and to discuss its opportunities for and threats to education. The strengths include using a sophisticated natural language model to generate plausible answers, self-improving capability, and providing personalised and real-time responses. As such, ChatGPT can increase access to information, facilitate personalised and complex learning, and decrease teaching workload, thereby making key processes and tasks more efficient. The weaknesses are a lack of deep understanding, difficulty in evaluating the quality of responses, a risk of bias and discrimination, and a lack of higher-order thinking skills. Threats to education include a lack of understanding of the context, threatening academic integrity, perpetuating discrimination in education, democratising plagiarism, and declining high-order cognitive skills. We provide agenda for educational practice and research in times of ChatGPT.

Title: "How to Harness Generative AI to Accelerate Human Learning"
Author: Johnson
Source: [46]
Score: ★★★★☆☆★☆☆ (6)
Annotations: Paper on the upsides and the potential of AI in learning (and therefor, education).
Abstract: The advent of generative AI has caused both excitement and anxiety among educators. Some school systems have gone so far as to ban it altogether. Generative AI has the potential to transform human learning; but like any new technology, it has both strengths and weaknesses, and adopting it involves some risks. There are risks that generative AI will mislead learners with wrong information, or that learners will use it to do their homework and take tests for them. This article presents some ways to take best advantage of generative AI, while managing and mitigating the risks. It also suggests some uses of generative AI to avoid. These insights are informed by learning science and extensive experience developing AI-enabled learning products. If applied properly, generative AI can dramatically accelerate human learning and do so at scale.

Title: "Students' Voices on Generative AI: Perceptions, Benefits, and Challenges in Higher Education"
Author: Chan and Hu
Source: [16]
Score: ★★★★★★★★★★ (10)
Annotations: Important paper on studetns perception and use of gen AI, taking away many unsupported assumptions on students though and behaviour relating to gen AI tools.
Abstract: This study explores university students' perceptions of generative AI (GenAI) technologies, such as ChatGPT, in higher education, focusing on familiarity, their willingness to engage, potential benefits and challenges, and effective integration. A survey of 399 undergraduate and postgraduate students from various disciplines in Hong Kong revealed a generally positive attitude towards GenAI in teaching and learning. Students recognized the potential for personalized learning support, writing and brainstorming assistance, and research and analysis capabilities. However, concerns about accuracy, privacy, ethical issues, and the impact on personal development, career prospects, and societal values were also expressed. According to John Biggs' 3P model, student perceptions significantly influence learning approaches and outcomes. By understanding students' perceptions, educators and policymakers can tailor GenAI technologies to address needs and concerns while promoting effective learning outcomes. Insights from this study can inform policy development around the integration of GenAI technologies into higher education. By understanding students' perceptions and addressing their concerns, policymakers can create well-informed guidelines and strategies for the responsible and effective implementation of GenAI

tools, ultimately enhancing teaching and learning experiences in higher education.

Title: “Generative Artificial Intelligence in Higher Education: Why the ’Banning Approach’ to Student Use is Sometimes Morally Justified”
Author: Fine Licht
Source: [103]
Score: ★★★★☆☆★★☆☆ (6)
Annotations: When the banning approach is or isn’t justified, interesting concept and impacting many assumptions.
Abstract: There has been significant discussion among academics and policymakers about managing the use of generative artificial intelligence tools, such as ChatGPT, Gem- ini, and GitHub Copilot, in higher education, particularly regarding student usage (Eke, 2023, malmstromg et al., 2023, Yeadon et al. 2023). Many universities have adopted a ’nuanced approach,’ which encourages responsible use of these tools to achieve high-quality outcomes while adhering to ethical principles and regula- tions (McDonald et al., 2024). This middle-ground stance is recommended by many involved in the policy debate (Gimpel et al., 2023; Rudolph et al., 2023; Slimi and Carballido, 2023). The belief is that these tools are largely beneficial, that their neg- ative effects can be managed, and that they will become ubiquitous, making resist- ance futile. Not all institutions align with this approach—some are more welcoming of these technologies, even mandating their use, while others ban or strongly dis- courage their use (McDonald et al., 2024).

Title: “When Banning Isn’t an Option: Embracing AI in Requirement Engineering Education”
Author: Moravánszky
Source: [92]
Score: ★★☆☆☆★★★☆☆ (5)
Annotations: View on implementing AI in education instead of banning it.
Abstract: In the dynamic landscape of education, the integration of Artificial Intelligence (AI) has become increasingly apparent. This experience report explores how students can be guided to use AI in the context of writing requirements, all within a bachelor module on Requirements Engineering at a Higher Education Institute (HEI). The study subject was a student assignment, graded using an evaluation rubric. Further insights were gathered through a questionnaire. S

Title: “A Comprehensive AI Policy Education Framework for University Teaching and Learning”
Author: Chan
Source: [104]
Score: ★★★★★★★★★★ (10)
Annotations: Publication of an early AI policy framework for higher education, in the beginning of the chatGPT hype.
Abstract: This study aims to develop an AI education policy for higher education by examining the perceptions and implications of text generative AI technologies. Data was collected from 457 students and 180 teachers and staff across various disciplines in Hong Kong universities, using both quantitative and qualitative research methods. Based on the findings, the study proposes an AI Ecological Education Policy Framework to address the multifaceted implications of AI integration in university teaching and learning. This framework is organized into three dimensions: Pedagogical, Governance, and Operational. The Pedagogical dimension concentrates on using AI to improve teaching and learning outcomes, while the Governance dimension tackles issues related to privacy, security, and accountability. The Operational dimension addresses matters concerning infrastructure and training. The framework fosters a nuanced understanding of the implications of AI integration in academic settings, ensuring that stakeholders are aware of their responsibilities and can take appropriate actions accordingly.

Title: “Prompting Change: Exploring Prompt Engineering in Large Language Model AI and Its Potential to Transform Education”
Author: Cain
Source: [93]
Score: ★★★★★★☆★☆☆ (7)
Annotations: Prompt engineering can be an important component of future education, teaching students how to use AI tools effectively, especially LLM tools. This papers zooms in on this topics and explores the potential of using prompt engineering and LLMs in education.
Abstract: This paper explores the transformative potential of Large Language Models Artificial Intelligence (LLM AI) in educational contexts, particularly focusing on the innovative practice of prompt engineering. Prompt engineering, characterized by three essential components of content knowledge, critical thinking, and iterative design, emerges as a key mechanism to access the transformative capabilities of LLM AI in the learning process. This paper charts the evolving trajectory of LLM AI as a tool poised to reshape educational practices and assumptions. In particular, this paper breaks down the potential of prompt engineer-

ing practices to enhance learning by fostering personalized, engaging, and equitable educational experiences. The paper underscores how the natural language capabilities of LLM AI tools can help students and educators transition from passive recipients to active co-creators of their learning experiences. Critical thinking skills, particularly information literacy, media literacy, and digital citizenship, are identified as crucial for using LLM AI tools effectively and responsibly. Looking forward, the paper advocates for continued research to validate the benefits of prompt engineering practices across diverse learning contexts while simultaneously promoting potential defects, biases, and ethical concerns related to LLM AI use in education. It calls upon practitioners to explore and train educational stakeholders in best practices around prompt engineering for LLM AI, fostering progress towards a more engaging and equitable educational future.

Title: *The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity*
Author: Shojaee*† *et al.*
Source: [91]
Score: ★★★★★☆☆★★★ (8)
Annotations: Paper on the limitations of LLM technology, by examen it capabilities of complex problem solving by giving different levels of complex problems. Proofing it is a smart guessing machine, instead of actual intelligence.
Abstract: Recent generations of frontier language models have introduced Large Reasoning Models (LRMs) that generate detailed thinking processes before providing answers. While these models demonstrate improved performance on reasoning benchmarks, their fundamental capabilities, scaling properties, and limitations remain insufficiently understood. Current evaluations primarily focus on established mathematical and coding benchmarks, emphasizing final answer accuracy. However, this evaluation paradigm often suffers from data contamination and does not provide insights into the reasoning traces' structure and quality. In this work, we systematically investigate these gaps with the help of controllable puzzle environments that allow precise manipulation of compositional complexity while maintaining consistent logical structures. This setup enables the analysis of not only final answers but also the internal reasoning traces, offering insights into how LRMs "think". Through extensive experimentation across diverse puzzles, we show that frontier LRMs face a complete accuracy collapse beyond certain complexities. Moreover, they exhibit a counter-intuitive scaling limit: their reasoning effort increases with problem complexity up to a point, then declines despite having an adequate token budget. By comparing LRMs with their standard LLM counterparts under equivalent inference compute, we identify three performance regimes: (1) low-complexity tasks where standard models surprisingly outperform LRMs, (2) medium-complexity tasks where additional thinking in LRMs demonstrates advantage, and (3) high-complexity tasks where both models experience complete collapse. We found that LRMs have limitations in exact computation: they fail to use explicit algorithms and reason inconsistently across puzzles. We also investigate the reasoning traces in more depth, studying the patterns of explored solutions and analyzing the models' computational behavior, shedding light on their strengths, limitations, and ultimately raising crucial questions about their true reasoning capabilities.

10.2 Grey literature

Title: *200 LLM Benchmarks and Evaluation Datasets*
Source: [36]
Annotations: Large dataset on LLM performances and capabilities

Title: "LiveBench: A Challenging, Contamination-Free LLM Benchmark"
Source: [27]
Annotations: Contamination-Free LLM Benchmark containing a leaderbord and scores of major and smaller LLM'sw

Title: *AI Index Annual Report 2024*
Source: [26]
Annotations: Standfords yearly AI index: The mission of the AI Index is to provide unbiased, rigorously vetted, and globally sourced data for policymakers, researchers, journalists, executives, and the general public to develop a deeper understanding of the complex field of AI. To achieve this, we track, collate, distill, and visualize data relating to artificial intelligence.

Title: *AI Leap 2025: Estonia Sets AI Standard in Education*
Source: [65]
Annotations: Estonia website on their AI integration in education initiative.

Title: *AI in Education: Establishing Foundations for Personalised Learning*
Source: [64]
Annotations: Estonia website on their AI integration in education initiative.

Title: *Estonia: AI Leap Initiative to Enhance Learning and Teaching*
Source: [63]
Annotations: Europe on Estonia's initiatives on AI integration into education.

Title: *Estonia's Schools Are Using ChatGPT to Personalize Education*
Source: [62]
Annotations: OpenAI on their collaboration with the Estonian government on AI integration in education.

Title: *Artificial intelligence and education and skills*
Annotations: OECD on Artificial intelligence and education and skills

Title: *Putting AI to the test*
Source: [105]
Annotations: OECD on current capabilities of AI systems (in education)

Title: *Is Education Losing the Race with Technology*
Source: [58]
Annotations: OECD on educational practices and AI capabilities

Title: *Impact of Advanced AI on Society*
Source: [68]
Annotations: Paper by Japan's MEXT department on the Impact of Advanced AI on Society, also elaborating on the AI integration in education.

Title: *Use of Generative AI: Guidelines for the Use of Generative AI in Primary and Secondary Education (Ver. 2.0)*
Source: [66]
Annotations: The webpage at 16-6-2025 on Use of generative AI, created by the Ministry of Education, Culture, Sports, Science and Technology. Based on research and pilot programs.

Title: *Guidelines for the Use of Generative AI in Primary and Secondary Education*
Source: [67]
Annotations: Guidelines for the Use of Generative AI in Primary and Secondary Education Ministry of Education, Culture and Science Directorate-General for Primary and Secondary Education 6th Year of Reiwa December 26 Ver. 2.0. This original document is in Japaneese, but is translated for this research. The original is included as a source.

Title: *White Paper on Science, Technology, and Innovation 2024 (Provisional Transla-*

tion): How AI will Transform Science, Technology and Innovation
Source: [87]
Annotations: Outline from 2024, White Paper on Science, Technology, and Innovation. Also mentioning the AI developments.

Title: *Derek Li And Squirrel Ai Aim To Lead The Future Of AI-Driven Education*
Source: [71]
Annotations: Forbes article on Squirrel AI, the background and it's impact.

Title: *Squirrel Ai Learning*
Source: [70]
Annotations: Short page by the World Economic Form on Squirrel AI

Title: *The Future of Learning is Here*
Source: [69]
Annotations: Homepage of SquirrelAI in July 2025, including some claims on numbers of students, effectiveness and world wide implementation.

Title: "When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction"
Source: [43]
Annotations: Study in effective adaptive learning, including squirrel Ai in it's comparison.

Title: *AI helps create personalized math lessons for students*
Source: [44]
Annotations: Microsoft article on collaboration, research and AI use with Eedi

Title: *Eedi's approach to AI*
Source: [73]
Annotations: Eedi AI policies and practices, from their own blog

Title: *Help all your students succeed in maths*
Source: [72]
Annotations: Homepage including numbers on students, schools and more

Title: "AI in Education and Learning Analytics in Singapore: An Overview of Key Projects and Initiatives"
Source: [106]
Annotations: Study on Ai in education initiatives in Singapore

Title: *Artificial Intelligence in Education*
Source: [107]
Annotations: Singapore's AI in education statement

Title: *AI in Education: Transforming Singapore's education system with student learning space*
Source: [74]
Annotations: Singapore's AI in education statement extended

Title: *Accelerating AI in Singapore: Develop Generational AI Capability Programmes to Build AI Aware and AI Ready Talents*
Source: [75]
Annotations: Page on AI4E in singapore

Title: *What You Need to Know About UNESCO's New AI Competency Frameworks for Students and Teachers*
Source: [25]
Annotations: To help education systems keep pace, UNESCO is launching two new AI competency frameworks - one for students and one for teachers.

Title: *AI Competency Framework for Students*
Source: [56]
Annotations: new AI competency frameworks for students

Title: *AI Competency Framework for Teachers*
Source: [57]
Annotations: new AI competency framework for teachers

Title: *AI for Teaching and Learning*
Source: [59]
Annotations: European School Education Platform stand on AI for teaching and learning

Title: *Adapting to the Future: Responsibly Integrating AI into Teaching and Learning*
Source: [61]
Annotations: European School Education Platform on how to responsibly integrating AI into teaching and learning

Title: *Enhancing Teaching and Learning Through Artificial Intelligence*
Source: [60]
Annotations: More on how artificial intelligence is used to help learners and teachers at the vocational school in Luxembourg.

Title: *AI Guidance & FAQs*
Source: [31]
Annotations: Harvard AI policies

Title: *Guidelines for Using ChatGPT and Other Generative AI Tools at Harvard*
Source: [76]
Annotations: Harvard's AI guidelines

Title: *New Guidelines for the Use of Generative AI in Education*
Source: [79]
Annotations: ETH AI guidelines

Title: *Academic Integrity*
Source: [80]
Annotations: ETH AI policies

Title: *PKUSTL Generative AI Rules for Inclusion in Academic Integrity Code*
Source: [34]
Annotations: Pekining University AI statement

Title: *Center for Artificial Intelligence and Machine Learning – CIAAM – Research Center at USP*
Source: [77]
Annotations: CIAAM webpage, university department for research on AI

Title: "A Proposal in Brazil to Use Generative AI in Education Threatens Quality and Equity"
Source: [78]
Annotations: Paper on concerns on AI proposal in Brasil

Title: *Interim SU Guidelines on Allowable AI Use and Academic Integrity*
Source: [94]
Annotations: Draft guidelines on how AI is allowed at the university of Stellenbosh

Title: *Position Statement on the Ethical Use of AI in Research and Teaching-Learning-Assessment*
Source: [32]
Annotations: Stellenbosh position statement on AI use in education

Title: *Uva AI CHAT*
Source: [81]
Annotations: UvA's statement on their inhouse AI tool for teachers and students, that will go live at the end of 2025. Mainly for data integrity.

Title: *Onze onderzoeksthema's*
Source: [83]
Annotations: Article on research focus relating to AI by the TU Delft

Title: *AI Onderwijsinnovatie*
Source: [82]
Annotations: Initiatives of the TU Delft on AI integration in Education

Title: *AI Onderwijs*
Source: [33]
Annotations: AI in Education at the TU Delft

Title: *NOLAI | Nationaal Groeifonds*
Source: [84]
Annotations: Rijksoverheid on the dutch NOLAI initiative, integrating AI in education and sociality. The project runs from 2022 to 2023.

Title: *Kwaliteitscriteria van toetsing*
Source: [55]

Annotations: Radboud university on quality assurance of assessment.

Title: *Kwaliteitsborging van toetsing*
Source: [54]
Annotations: UvA on quality assurance of assessment.

Title: *Policy Framework and Guidelines on GenAI in Education*
Source: [108]
Annotations: Uva's updated framework for Ai policies and guidelines

Title: *Assessment Policy Addendum: Focus on GenAI in Assessment*
Source: [109]
Annotations: Uva's updated assesment methods for Ai policies and guidelines

Title: *Beleidskader GenAI in het Onderwijs*
Source: [95]
Annotations: Uva's updated document on Ai policies and guidelines

10.3 Tertiary sources

Title: *Cambridge Dictionary Online*
Source: [5]
Annotations: English dictionary used to define the exact meaning of words.

Title: *University of Amsterdam*
Source: [1]
Annotations: General U.v.A. website, for the use of quotes and factual information about the uva.

Title: *VU-UvA Task Force on AI in Education*
Source: [15]
Annotations: More information on the U.v.A. and V.U. actions on A.I. policies in education.

Title: *AI policy*
Source: [12]
Annotations: More information on current LLM use in education at the U.v.A.

Title: *Plagiaat en fraude*
Source: [14]
Annotations: More information on plagiarism at the U.v.A.

Title: *AI-tools en je studie*
Source: [13]
Annotations: More information on the use of A.I. tools at the U.v.A.

Title: *Curriculum*
Source: [21]
Annotations: Definition of curriculum by UNESCO Institute for Statistics

Bibliography

[1] University of Amsterdam. "University of amsterdam." Accessed: 2025-04-22. (2025), [Online]. Available: https://www.uva.nl.

[2] L. Labadze, M. Grigolia, and L. Machaidze, "Role of ai chatbots in education: Systematic literature review," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 56, 2023. DOI: https://doi.org/10.1186/s41239-023-00422-0. [Online]. Available: https://doi.org/10.1186/s41239-023-00422-0.

[3] C. K. Lo, "What is the impact of chatgpt on education? a rapid review of the literature," *Educ. Sci.*, vol. 13, no. 4, p. 410, 2023. DOI: https://doi.org/10.3390/educsci13040410.

[4] H. Wang, A. Dang, Z. Wu, and S. Mac, "Generative ai in higher education: Seeing chatgpt through universities' policies, resources, and guidelines," *Computers and Education: Artificial Intelligence*, vol. 5, p. 100 326, 2024. [Online]. Available: https://doi.org/10.1016/j.caeai.2024.100326.

[5] Cambridge University Press. "Cambridge dictionary online." Accessed: 2025-04-11. (2024), [Online]. Available: https://dictionary.cambridge.org/.

[6] A. Bhardwaj, "Importance of education in human life: A holistic approach," *International Journal of Science and Consciousness*, vol. 2, no. 2, pp. 23–28, Jun. 2016. [Online]. Available: https://www.researchgate.net/publication/366005991_Pages_23-28_eISSN_2455-2038_An_international_refereed_quarterly_open_access_multidisciplinary_e-journal_Overview_Article_International_Journal_of_Science_and_Consciousness_Access_online_at_wwwijsc.

[7] J. W. Meyer, "The effects of education as an institution," *American Journal of Sociology*, vol. 83, no. 1, pp. 55–77, Jul. 1977. DOI: 10.1086/226506. [Online]. Available: https://doi.org/10.1086/226506.

[8] P. W. Kingston, R. Hubbard, B. Lapp, P. Schroeder, and J. Wilson, "Why education matters," *Sociology of Education*, vol. 76, no. 1, pp. 53–70, Jan. 2003. DOI: 10.2307/3090261. [Online]. Available: https://www.jstor.org/stable/3090261.

[9] R. Dale and S. Robertson, "Editorial introduction," *Globalisation, Societies and Education*, vol. 1, no. 1, pp. 3–11, 2003, Accessed via Crossref. DOI: 10.1080/1476772032000060543.

[10] OECD, "Education at a glance 2024: Oecd indicators," 2024. [Online]. Available: https://doi.org/10.1787/c00cad36-en.

[11] Y. Albadarin, M. Saqr, N. Pope, and M. Tukiainen, "A systematic literature review of empirical research on chatgpt in education," *Discover Education*, vol. 3, no. 60, 2024. [Online]. Available: https://doi.org/10.1007/s44217-024-00084-x.

[12] U. van Amsterdam, *Ai policy*, https://www.uva.nl/en/about-the-uva/about-the-university/ai/ai-policy/ai-policy.html, Accessed: 2025-04-16.

[13] U. van Amsterdam, *Ai-tools en je studie*, https://student.uva.nl/onderwerpen/ai-tools-en-je-studie, Accessed: 2025-04-16.

[14] U. van Amsterdam, *Plagiaat en fraude*, https://student.uva.nl/onderwerpen/plagiaat-en-fraude, Accessed: 2025-04-16.

[15] U. van Amsterdam, *Vu-uva task force on ai in education*, https://www.uva.nl/en/about-the-uva/about-the-university/ai/ai-in-education/vu-uva-task-force-on-ai-in-education/vu-uva-task-force-on-ai-in-education.html, Accessed: 2025-04-16.

[16] C. K. Y. Chan and W. Hu, "Students' voices on generative ai: Perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, Article number 43, 2023, Published: 2023. DOI: 10.1186/s41239-023-00411-8. [Online]. Available:

https://doi.org/10.1186/s41239-023-00411-8.

[17] J. Qadir, "Engineering education in the era of chatgpt: Promise and pitfalls of generative ai for education," in *2023 IEEE Global Engineering Education Conference (EDUCON)*, Baku, Azerbaijan: IEEE, 2023, pp. 1–9. DOI: 10.1109/EDUCON53801.2023.10109764. [Online]. Available: https://ieeexplore.ieee.org/document/10109764.

[18] UNESCO, F. Miao, W. Holmes, R. Huang, and H. Zhang, *AI and education: guidance for policy-makers*. UNESCO, 2021, p. 45, ISBN: 978-92-3-100447-6. DOI: https://doi.org/10.54675/PCSP7350.

[19] J. Mander, *Four Arguments for the Elimination of Television*. New York: William Morrow and Company, 1978, Accessed: 2025-01-10. [Online]. Available: https://slinkysearch.com.au.

[20] N. Carr, *The Shallows: How the Internet Is Changing the Way We Think, Read and Remember*. New York: W. W. Norton & Company, 2010, Accessed: 2025-01-10. [Online]. Available: https://books.google.com.

[21] U. I. for Statistics, *Curriculum*, Accessed: 2024-07-15, 2024. [Online]. Available: https://uis.unesco.org/node/3297441.

[22] J. Hattie and E. M. Anderman, *Visible Learning Guide to Student Achievement: Schools Edition*, 1st. London: Routledge, 2019, eBook Published: 28 October 2019, ISBN: 9781351257848. DOI: 10.4324/9781351257848.

[23] M. J. Prince and R. M. Felder, "Inductive teaching and learning methods: Definitions, comparisons, and research bases," *Journal of Engineering Education*, vol. 95, no. 2, pp. 123–138, 2006. DOI: 10.1002/j.2168-9830.2006.tb00884.x.

[24] J. Hattie, *Visible Learning: The Sequel: A Synthesis of Over 2,100 Meta-Analyses Relating to Achievement*, 1st. London: Routledge, 2023, eBook Published: 20 March 2023, ISBN: 9781003380542. DOI: 10.4324/9781003380542.

[25] UNESCO, *What you need to know about unesco's new ai competency frameworks for students and teachers*, https://www.unesco.org/en/articles/what-you-need-know-about-unescos-new-ai-competency-frameworks-students-and-teachers?hub=32618, Published: 3 Sep 2024. Last updated: 20 Feb 2025. Accessed: 17 Jun 2025, 2025.

[26] S. I. for Human-Centered Artificial Intelligence (HAI), "Ai index annual report 2024," 2024. [Online]. Available: https://aiindex.stanford.edu/report/.

[27] C. White *et al.*, "Livebench: A challenging, contamination-free LLM benchmark," in *The Thirteenth International Conference on Learning Representations*, 2025. [Online]. Available: https://livebench.ai/#/.

[28] J. Qadir, "Engineering education in the era of chatgpt: Promise and pitfalls of generative ai for education," *2023 IEEE Global Engineering Education Conference (EDUCON)*, pp. 1–9, 2023. DOI: 10.1109/EDUCON53801.2023.10109764. [Online]. Available: https://ieeexplore.ieee.org/document/10109764.

[29] D. T. T. Mai, C. V. Da, and N. V. Hanh, "The use of chatgpt in teaching and learning: A systematic review through swot analysis approach," *Frontiers in Education*, 2024. DOI: https://doi.org/10.3389/feduc.2024.1354879. [Online]. Available: https://www.frontiersin.org/articles/10.3389/feduc.2024.1354879/full.

[30] S. Beltozar-Clemente, E. Díaz-Vega, J. Zapata-Paulini, and R. E. Tejeda-Navarrete, "We can rely on chatgpt as an educational tutor: A cross-sectional study of its performance, accuracy, and limitations in university admission tests," *International Journal of Engineering Pedagogy (iJEP)*, vol. 14, no. 1, pp. 50–60, 2024. DOI: 10.3991/ijep.v14i1.46787. [Online]. Available: https://online-journals.org/index.php/i-jep/article/view/46787.

[31] Harvard Faculty of Arts and Sciences, *Ai guidance & faqs*, https://oue.fas.harvard.edu/ai-guidance, Webpage. Published: July 2025. Accessed: 17 Jun 2025, 2025.

[32] Stellenbosch University, *Position statement on the ethical use of ai in research and teaching-learning-assessment*, Approved by SU Senate. Date of Approval: 22 November 2024. Implementation: 1 January 2025, 2024.

[33] TU Delft, *Ai onderwijs*, https://www.tudelft.nl/ai/onderwijs, Webpage article, category: General AI education approach, 2025.

[34] Peking University School of Transnational Law, *Pkustl generative ai rules for inclusion in academic integrity code*, Document. Fall 2024. STL's current AI policy in its Academic Integrity Code, as approved by the

faculty in 2024, 2024. [Online]. Available: https://stl.pku.edu.cn/uploads/202411/26102211_98533_67921156%5B-%5DSTLAcademicRules-AIPolicy-Fall2024.pdf.

[35] D. Kennedy, *Writing and Using Learning Outcomes: A Practical Guide*. Cork: University College Cork, 2006, Accessed via cora.ucc.ie. Widely cited in discussions of student-centered learning and curriculum design in the context of the Bologna Process. [Online]. Available: https://cora.ucc.ie/handle/10468/1613.

[36] Evidently AI, *200 llm benchmarks and evaluation datasets*, https://www.evidentlyai.com/llm-evaluation-benchmarks-datasets, A database of LLM benchmarks and datasets to evaluate the performance of language models. First published: December 10, 2024., Dec. 2024. [Online]. Available: https://www.evidentlyai.com/llm-evaluation-benchmarks-datasets.

[37] OECD. "Artificial intelligence and education and skills." (2025), [Online]. Available: https://www.oecd.org/en/topics/artificial-intelligence-and-education-and-skills.html (visited on 06/10/2025).

[38] J. D. Bransford, A. L. Brown, and R. R. Cocking, *How People Learn: Brain, Mind, Experience, and School*. Washington, D.C.: National Academy Press, 2000, Accessed via csun.edu. [Online]. Available: https://www.csun.edu/~SB4310/How%20People%20Learn.pdf.

[39] R. Parker, B. S. Thomsen, and A. Berry, "Learning through play at school – a framework for policy and practice," *Frontiers in Education*, vol. 7, Feb. 2022. DOI: 10.3389/feduc.2022.751801.

[40] M. Nilsson, B. Ferholt, and R. Lecusay, "'the playing-exploring child': Reconceptualizing the relationship between play and learning in early childhood education," *Contemporary Issues in Early Childhood*, vol. 19, no. 3, pp. 231–245, 2017, (Original work published 2018). DOI: 10.1177/1463949117710800.

[41] J. M. Zosh *et al.*, "Learning through play: A review of the evidence," *Danish University Colleges*, Nov. 2017, ResearchGate. PDF Available. DOI: 10.13140/RG.2.2.16823.01447.

[42] C. E. Hmelo-Silver, "Problem-based learning: What and how do students learn?" *Educational Psychology Review*, vol. 16, no. 3, pp. 235–266, Aug. 2004, Access provided by SURF B.V. DOI: 10.1023/B:EDPR.0000034022.16470.f3. [Online]. Available: https://link.springer.com/article/10.1023/B:EDPR.0000034022.16470.f3.

[43] S. Wang *et al.*, "When adaptive learning is effective learning: Comparison of an adaptive learning system to teacher-led instruction," *Interactive Learning Environments*, vol. 31, pp. 1–11, Aug. 2020. DOI: 10.1080/10494820.2020.1808794.

[44] J. Roach. "Ai helps create personalized math lessons for students." Accessed: 2025-06-17, Microsoft. (Jul. 13, 2022), [Online]. Available: https://news.microsoft.com/source/features/ai/eedi-online-math-quiz/.

[45] W. Holmes and I. Tuomi, "State of the art and practice in ai in education," *European Journal of Education*, vol. 57, no. 4, pp. 532–553, 2022, First published: 30 October 2022. DOI: 10.1111/ejed.12533. [Online]. Available: https://doi.org/10.1111/ejed.12533.

[46] W. L. Johnson, "How to harness generative ai to accelerate human learning," *International Journal of Artificial Intelligence in Education*, vol. 34, pp. 1287–1291, Sep. 2024, Published: 08 August 2023. DOI: 10.1007/s40593-023-00367-w. [Online]. Available: https://doi.org/10.1007/s40593-023-00367-w.

[47] G. D. Kuh *et al.*, *Using Evidence of Student Learning to Improve Higher Education*. Jossey-Bass, 2015. [Online]. Available: https://books.google.nl/books?hl=nl&lr=&id=mTitBQAAQBAJ&oi=fnd&pg=PR9&dq=how+can+higher+education+be+improved&ots=4k4xnzRTe7&sig=-qgdLYolKqvrLXBPMHHWqn9vcdM&redir_esc=y#v=onepage&q=how%20can%20higher%20education%20be%20improved&f=false.

[48] J. Spring, "Research on globalization and education," *Review of Educational Research*, vol. 78, no. 2, pp. 330–363, 2008, First published online June 1, 2008. DOI: 10.3102/0034654308317846. [Online]. Available: https://journals.sagepub.com/doi/10.3102/0034654308317846.

[49] R. Huang, "Comparative analysis of the differences in the education system in china vs germany," *Journal of Education*, Jun. 2024, Corpus ID: 270598389, Accessed via Semantic Scholar. DOI: 10.54097/3be6kt31. [Online]. Available: https://www.semanticscholar.org/paper/72b3e4b0d59f1df70ef29edd9ed51a416ed5bedf.

[50] J. B. Biggs, C. S.-k. Tang, and G. Kennedy, *Teaching for Quality Learning at University*, Fifth. Maidenhead: Open University Press, 2022.

[51] O. L. Liu, "Measuring learning outcomes in higher education," *R & D Connections*, no. 10, Aug. 2009. [Online]. Available: https://www.ets.org/Media/Research/pdf/RD_Connections10.pdf#:~:text=Under%20former%20U,organizations%20in%20higher%20education%2C%20the.

[52] D. Reynolds, P. Sammons, B. De Fraine, J. Van Damme, T. Townsend, and C. Teddlie, "Educational effectiveness research (eer): A state-of-the-art review," *School Effectiveness and School Improvement*, vol. 25, no. 2, pp. 197–230, 2014. DOI: 10.1080/09243453.2014.885450. [Online]. Available: https://doi.org/10.1080/09243453.2014.885450.

[53] N. Maassen, D. den Otter, S. Wools, B. Hemker, G. Straetmans, and T. Eggen, "Kwaliteit van toetsen binnen handbereik: Een reviewstudie van onderzoek en onderzoeksresultaten naar de kwaliteit van toetsen," RCEC, Cito, Arnhem, Tech. Rep., 2014. [Online]. Available: https://www.rcec.nl/wp-content/uploads/2019/04/Reviewstudie_KvTbH.pdf.

[54] H. Kousbroek and I. Pancras, *Kwaliteitsborging van toetsing*, Workshopmateriaal voor examencommissies over toetskwaliteit, 2025. [Online]. Available: https://tlc.uva.nl/article-category/toetsing/?faculty=55.

[55] Radboud Universiteit, *Kwaliteitscriteria van toetsing*, Geraadpleegd op 16 juli 2025, 2025. [Online]. Available: https://www.ru.nl/medewerkers/docenten/toetsen-en-beoordelen/basisontwerp-van-toetsing/kwaliteitscriteria.

[56] F. Miao, K. Shiohira, and N. Lao, *AI Competency Framework for Students*, English. UNESCO, 2024, p. 80, License: CC BY-SA 3.0 IGO. Available in: English, Français, Tiếng Việt, Español, ISBN: 978-92-3-100709-5. DOI: 10.54675/JKJB9835.

[57] F. Miao and M. Cukurova, *AI Competency Framework for Teachers*, English. UNESCO, 2024, p. 52, License: CC BY-SA 3.0 IGO. Available in: English, Français, Tiếng Việt, Español, ISBN: 978-92-3-100707-1. DOI: 10.54675/ZJTE2084.

[58] OECD. "Is education losing the race with technology." (2025), [Online]. Available: https://www.oecd.org/en/publications/is-education-losing-the-race-with-technology_73105f99-en.html (visited on 06/10/2025).

[59] European School Education Platform, *Ai for teaching and learning*, https://school-education.ec.europa.eu/en/discover/news/ai-teaching-and-learning, News item. Published: 27 May 2024. Accessed: 17 Jun 2025, 2024.

[60] European School Education Platform, *Enhancing teaching and learning through artificial intelligence*, https://school-education.ec.europa.eu/en/discover/practice/enhancing-teaching-and-learning-through-artificial-intelligence, Practice video. Published: 17 May 2024. Accessed: 17 Jun 2025, 2024.

[61] European School Education Platform, *Adapting to the future: Responsibly integrating ai into teaching and learning*, https://school-education.ec.europa.eu/en/discover/practice/adapting-future-responsibly-integrating-ai-teaching-and-learning, Practice article. Published: 17 Apr 2024. Featured item. Accessed: 17 Jun 2025, 2024.

[62] OpenAI. "Estonia's schools are using chatgpt to personalize education." Accessed: 2025-06-09. (2025), [Online]. Available: https://openai.com/index/estonia-schools-and-chatgpt/.

[63] Eurydice. "Estonia: Ai leap initiative to enhance learning and teaching." Accessed: 2025-06-09. (2025), [Online]. Available: https://eurydice.eacea.ec.europa.eu/news/estonia-ai-leap-initiative-enhance-learning-and-teaching.

[64] Education Estonia. "Ai in education: Establishing foundations for personalised learning." Accessed: 2025-06-09. (2024), [Online]. Available: https://www.educationestonia.org/ai-in-education-establishing-foundations-for-personalised-learning/.

[65] e-Estonia. "Ai leap 2025: Estonia sets ai standard in education." Accessed: 2025-06-09. (2025), [Online]. Available: https://e-estonia.com/ai-leap-2025-estonia-sets-ai-standard-in-education/.

[66] Ministry of Education, Culture, Sports, Science and Technology – Japan (MEXT), *Use of generative ai: Guidelines for the use of generative ai in primary and secondary education (ver. 2.0)*, Published December 26, 2024, 2024. [Online]. Available: https://www.mext.go.jp/a_menu/other/mext_02412.html.

[67] Ministry of Education, Culture, Sports, Science and Technology – Japan (MEXT), *Guidelines for the use of generative ai in primary and secondary education*, Directorate-General for Primary and Secondary Education, Ver. 2.0, December 26, 6th Year of Reiwa, 2024. [Online]. Available: https://www.mext.go.jp/content/20241226-mxt_shuukyo02-000030823_001.pdf.

[68] Ministry of Education, Culture, Sports, Science and Technology – Japan (MEXT), *Impact of advanced ai on society*, Policy Paper, 2024. [Online]. Available: https://www.mext.go.jp/en/content/20241224-mxt_chousei01-000036407-09.pdf.

[69] Squirrel AI. "The future of learning is here." Squirrel AI homepage. (2025), [Online]. Available: https://squirrelai.com/ (visited on 06/17/2025).

[70] World Economic Forum. "Squirrel ai learning." World Economic Forum Website. (2025), [Online]. Available: https://www.weforum.org/organizations/squirrel-ai-learning/ (visited on 06/17/2025).

[71] R. Flannery. "Derek li and squirrel ai aim to lead the future of ai-driven education." Forbes China. (Feb. 2025), [Online]. Available: https://www.forbes.com/sites/forbeschina/2025/02/18/derek-li-and-squirrel-ai-aim-to-lead-the-future-of-ai-driven-education/ (visited on 06/17/2025).

[72] "Help all your students succeed in maths." Accessed: 2025-06-17, Eedi. (Jul. 2025), [Online]. Available: https://eedi.com/.

[73] B. Groot. "Eedi's approach to ai." Accessed: 2025-06-17, Eedi. (Oct. 21, 2024), [Online]. Available: https://eedi.com/us/research-blog/eedis-approach-to-ai.

[74] Government Technology Agency of Singapore, *Ai in education: Transforming singapore's education system with student learning space*, https://www.tech.gov.sg/technews/ai-in-education-transforming-singapore-education-system-with-student-learning-space, Accessed: 21 Jan 2025, 2025.

[75] AI Singapore, *Accelerating ai in singapore: Develop generational ai capability programmes to build ai aware and ai ready talents*, https://learn.aisingapore.org/home-2-2/, Accessed: 17 Jun 2025, 2025.

[76] Harvard University, *Guidelines for using chatgpt and other generative ai tools at harvard*, https://provost.harvard.edu/guidelines-using-chatgpt-and-other-generative-ai-tools-harvard, Webpage. Published: July 2025. Accessed: 17 Jun 2025, 2025.

[77] Center for Artificial Intelligence and Machine Learning (CIAAM), *Center for artificial intelligence and machine learning – ciaam – research center at usp*, https://ciaam.usp.br/en/, Website. Accessed: 17 Jun 2025, 2025.

[78] F. D. Matos, G. Girotto Junior, A. d. M. Arnt, and A. Lippi, "A proposal in brazil to use generative ai in education threatens quality and equity," *Patterns*, vol. 5, no. 7, p. 101 020, 2024, ISSN: 2666-3899. DOI: 10.1016/j.patter.2024.101020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666389924001533.

[79] ETH Zurich, *New guidelines for the use of generative ai in education*, https://ethz.ch/staffnet/en/news-and-events/internal-news/archive/2024/07/new-guidelines-for-the-use-of-generative-ai-in-education.html, Web article. Published: July 2025. Accessed: 17 Jun 2025, 2025.

[80] ETH Zurich, *Academic integrity*, https://ethz.ch/en/the-eth-zurich/education/ai-in-education/academic-integrity.html, Web article. Published: July 2025. Accessed: 17 Jun 2025, 2025.

[81] University of Amsterdam, *Uva AI CHAT*, https://www.uva.nl/en/about-the-uva/about-the-university/ai/ai-in-education/uva-ai-chat/uva-ai-chat.html, Webpage article, AI in education category, 2025.

[82] TU Delft, *Ai onderwijsinnovatie*, https://www.tudelft.nl/ai/onderwijs/ai-onderwijsinnovatie, Webpage article, category: AI in education innovation, 2025.

[83] TU Delft, *Onze onderzoeksthema's*, https://www.tudelft.nl/ai/onderzoek-innovatie/onze-onderzoeksthemas, Webpage article, category: AI research themes, 2025.

[84] Rijksoverheid, *Nolai | nationaal groeifonds*, https://www.nationaalgroeifonds.nl/overzicht-lopende-projecten/thema-onderwijs/nationaal-onderwijslab, Webpage article, category: National AI Education Lab (NOLAI), 2025.

[85] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, vol. 30, no. 1-7, pp. 107–117, 1998. DOI: 10.1016/S0169-7552(98)00110-X. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S016975529800110X.

[86] A. Zuiderwijk, Y.-C. Chen, and F. Salem, "Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda," *Government Information Quarterly*, vol. 38, no. 1, p. 101 577, 2021. DOI: 10.1016/j.giq.2021.101577. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0740624X21000137.

[87] Ministry of Education, Culture, Sports, Science and Technology – Japan (MEXT), *White paper on science, technology, and innovation 2024 (provisional translation): How ai will transform science, technology and innovation*, White Paper, 2024. [Online]. Available: https://www.mext.go.jp/en/content/20240611-ope_dev03-000036407-1.pdf#:~:text=%E2%80%A2%20MEXT%20has%20announced%20tentative,knowledge%20are%20now%20being%20accumulated.

[88] P. R. Burden and D. M. Byrd, *Methods for Effective Teaching: Meeting the Needs of All Students*, 8th ed. Hoboken, NJ: Pearson Education, Inc., 2019, Library of Congress Classification: LB1025.3 .B87 2019; DDC: 371.102–dc23. LC record available at https://lccn.loc.gov/2017040176, ISBN: 9780134695747.

[89] M. Farrokhnia, S. K. Banihashem, O. Noroozi, and A. Wals, "A swot analysis of chatgpt: Implications for educational practice and research," *Innovations in Education and Teaching International*, pp. 460–474, 2023. DOI: 10.1080/14703297.2023.2195846. [Online]. Available: https://doi.org/10.1080/14703297.2023.2195846.

[90] D. E. Knuth, *The Art of Computer Programming*, Hardcover. Engels, 1968, p. 736, ISBN: 9780137935109.

[91] P. Shojaee*†, I. Mirzadeh*, K. Alizadeh, M. Horton, S. Bengio, and M. Farajtabar, *The illusion of thinking: Understanding the strengths and limitations of reasoning models via the lens of problem complexity*, 2025. [Online]. Available: https://ml-site.cdn-apple.com/papers/the-illusion-of-thinking.pdf.

[92] A. Moravánszky, "When banning isn't an option: Embracing ai in requirement engineering education," *Unspecified Journal*, 2023.

[93] W. Cain, "Prompting change: Exploring prompt engineering in large language model ai and its potential to transform education," *Journal of Science Education and Technology*, vol. 68, pp. 47–57, Oct. 2023. DOI: 10.1007/s11528-023-00896-0. [Online]. Available: https://link.springer.com/article/10.1007/s11528-023-00896-0.

[94] Stellenbosch University, *Interim su guidelines on allowable ai use and academic integrity*, Internal policy document. Version: Draft, 2025. Focused on responsible use of AI in assessment, 2025.

[95] Universiteit van Amsterdam, *Beleidskader genai in het onderwijs*, Toegang verkregen op 17 juli 2025, May 2025. [Online]. Available: https://www.uva.nl/over-de-uva/over-de-universiteit/ai/ai-in-het-onderwijs/vu-uva-taskforce-ai-in-het-onderwijs/advies-vu-uva-taskforce-ai-in-het-onderwijs.html.

[96] E. Schofer, F. O. Ramirez, and J. W. Meyer, "The societal consequences of higher education," *Sociology of Education*, vol. 94, no. 1, 2020, First published online July 21, 2020. DOI: 10.1177/0038040720942912. [Online]. Available: https://doi.org/10.1177/0038040720942912.

[97] R. Dale, "Globalisation, knowledge economy and comparative education," *Comparative Education*, vol. 41, no. 2, pp. 117–149, 2005, Accessed via Crossref and Web of Science. DOI: 10.1080/03050060500150906.

[98] C. Thieme, V. Giménez, and D. Prior, "A comparative analysis of the efficiency of national education systems," *Asia Pacific Education Review*, vol. 13, pp. 1–15, 2012, Published online: 16 July 2011, Access provided by University of Amsterdam. DOI: 10.1007/s12564-011-9177-6. [Online]. Available: https://link.springer.com/article/10.1007/s12564-011-9177-6.

[99] S. Qalandarova and L. Iskandarova, "Education systems of countries in comparison," *International Multidisciplinary Journal for Research & Development*, vol. 11, no. 11, 2024, Retrieved from ijmrd.in. [Online]. Available: https://www.ijmrd.in/index.php/imjrd/article/view/2043.

[100] O. K. Kilag, J. P. Zarco, J. M. V. Sasan, *et al.*, "How does philippines's education system compare to finland's?" *Basic and Applied Education Research Journal*, Jun. 2023, Corpus ID: 259292207. DOI: 10.11594/baerj.04.01.03.

[101] J. Cowan, "Teaching for quality learning at university - by john biggs & catherine tang," *British Journal of Educational Technology*, vol. 43, no. 3, E94–E95, 2012, Book Review.

[102] K. Zhang and A. B. Aslan, "Ai technologies for education: Recent research & future directions," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100 025, 2021. DOI: 10.1016/j.caeai.2021.100025. [Online]. Available: https://doi.org/10.1016/j.caeai.2021.100025.

[103] K. de Fine Licht, "Generative artificial intelligence in higher education: Why the 'banning approach' to student use is sometimes morally justified," *Philosophy & Technology*, vol. 37, no. 113, Aug. 2024. DOI: 10.1007/s13347-024-00799-9.

[104] C. K. Y. Chan, "A comprehensive ai policy education framework for university teaching and learning," *International Journal of Educational Technology in Higher Education*, vol. 20, Article number 38, 2023, Published: 07 July 2023, Open access. DOI: 10.1186/s41239-023-00392-7. [Online]. Available: https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-023-00392-7.

[105] OECD. "Putting ai to the test." (2025), [Online]. Available: https://www.oecd.org/en/publications/putting-ai-to-the-test_2c297e0b-en.html (visited on 06/10/2025).

[106] A. Lee, E. Koh, and C.-K. Looi, "Ai in education and learning analytics in singapore: An overview of key projects and initiatives," *Information and Technology in Education and Learning*, vol. 3, Inv–p001, Nov. 2023. DOI: 10.12937/itel.3.1.Inv.p001. [Online]. Available: https://www.researchgate.net/publication/375137743_AI_in_Education_and_Learning_Analytics_in_Singapore_An_Overview_of_Key_Projects_and_Initiatives.

[107] Ministry of Education, Singapore, *Artificial intelligence in education*, https://www.moe.gov.sg/education-in-sg/educational-technology-journey/edtech-masterplan/artificial-intelligence-in-education, Last updated: 20 Sep 2023, 2023.

[108] University of Amsterdam, *Policy framework and guidelines on genai in education*, Accessed: 2025-07-17, May 2025. [Online]. Available: https://www.uva.nl/binaries/content/assets/uva/nl/over-de-uva/over-de-uva/beleid-en-financien/onderwijs/beleidskader_genai_eng_040625.pdf.

[109] University of Amsterdam, *Assessment policy addendum: Focus on genai in assessment*, Accessed: 2025-07-17, May 2025. [Online]. Available: https://www.uva.nl/binaries/content/assets/uva/nl/over-de-uva/over-de-uva/beleid-en-financien/onderwijs/addendum_toetsbeleid_eng_030625.pdf.

Discography

You might be wondering how we got here. Well, the truth is, this chapter reveals the secret to surviving a thesis: really good music. Not the kind you play during deep thinking, but the kind that keeps you going when you're writing (and struggling). Below are the albums and songs I listened to the most. May they inspire or comfort you as you write your own masterpiece.

Albums that got me through

- Hallucinating Love - Maribou State
- El bueno y el malo - Hermanos Gutierrez
- Circles - Mac Miller
- Flow State - Tash Sultana
- Het is een Wies - Wies

Artists that you should know

- The streets
- Franc Moody
- J. Bernardt
- Thijs Boontjes
- Fox Stevenson

Podcasts that enrich your life

- Hoe word je een podkast
- Dear Hank and John
- Sommar & Vinter
- De grote podcastlas
- Brainwash

Soundtracks that bring you to another world

- The Chronicles of Narnia: The Lion, The Witch and The Wardrobe (Original Motion Picture Soundtrack)
- Interstellar (Original Motion Picture Soundtrack)
- The Grand Budapest Hotel (Original Soundtrack)
- About Time (Soundtrack)
- A'dam: E.V.A. (Soundtrack)

Songs That hit just right

- Linde met een E - Lucky Fonz III
- Sleepwalking - Issey Cross, Songer
- Oeps - Sef, Froukje
- Loud Pipes - Ratatat
- Let's go back - Jungle

Songs you leave on repeat

- Song for Denise - Piano Fantasia
- I only smoke when I drink - Nimino
- Want to Love - Alobio
- Assumptions - Sam Gellaitry
- Birth4000 - Floating Points

Songs for when you are going full machine mode

- Create Machines - Venjent
- Prototype - Savant
- Working Machines - Tom Trago
- Christine - Joyhausser
- Galvanic - Bytheway-May
- Noting Ontploffing - Elmer
- The Magic Bomb - Hoang Read
- Miss you - Bad Smith
- Lavender City - PERMSKY KRAY
- End of Line - Daft Punk

Glossary

AI Artificial Intelligence. 1, 7, 8, 10–16, 18–20, 22, 24, 34–46, 48–53
AIED Artificial Intelligence in Education. 10, 11, 13, 17, 40, 53
GenAI Generative Artificial Intelligence. 8, 11–14, 16, 17, 20, 27, 35–40
LLM Large Language Model. 1, 7–17, 19–21, 24, 25, 27, 30, 33–35, 37, 38, 42, 43, 45, 46, 50, 52, 53
OECD The Organisation for Economic Co-operation and Development. 35
UvA University of Amsterdam. 7, 8, 14, 40

Appendix A

LiveBench LLM benchmark

5/13/25, 11:20 AM LiveBench

LiveBench

A Challenging, Contamination-Free LLM Benchmark

LiveBench will appear as a Spotlight Paper in ICLR 2025.
This work is sponsored by Abacus.AI

Leaderboard Details Code Data Paper

New! Check out LiveSWEBench, our new benchmark for AI coding agents

Introduction

Introducing **LiveBench**: a benchmark for LLMs designed with test set contamination and objective evaluation in mind. It has the following properties:

- LiveBench limits potential contamination by releasing new questions regularly.
- Each question has verifiable, objective ground-truth answers, eliminating the need for an LLM judge.
- LiveBench currently contains a set of 18 diverse tasks across 6 categories, and we will release new, harder tasks over time.

We will evaluate your model on LiveBench! Open a github issue or email us at livebench@livebench.ai!

Leaderboard

We update questions regularly so that the benchmark completely refreshes every 6 months. Some questions for previous releases are available here. The most recent version is **LiveBench-2025-04-25**. This verison includes updated coding and data analysis questions.

To further reduce contamination, we delay publicly releasing the questions from the most-recent updates. LiveBench-2025-04-02 and LiveBench-2025-04-25 comprise ~300 new questions, so currently ~30% of questions in LiveBench are not publicly released.

View Full Changelog

2025-04-25

- [x] Reasoning Average
- [ ] Show Subcategories
- [x] Coding Average
- [ ] Show Subcategories
- [x] Mathematics Average
- [ ] Show Subcategories
- [x] Data Analysis Average
- [ ] Show Subcategories
- [x] Language Average
- [ ] Show Subcategories
- [x] IF Average
- [ ] Show Subcategories

- [x] Show Organization
- [ ] Show API Name
- [x] Show Reasoning Models

Clear Filters

Search...

Filter by organization...

Model	Organization	Global Average	Reasoning Average	Coding Average	Mathematics Average	Data Analysis Average	Language Average	IF Average
o3 High	OpenAI	80.71	93.33	76.71	85.00	67.02	76.00	86.17
o3 Medium	OpenAI	79.25	91.00	77.86	80.66	68.19	73.48	84.32
Gemini 2.5 Pro Preview (2025-05-06)	Google	78.99	88.25	72.87	88.63	68.85	71.81	83.50
o4-Mini High	OpenAI	78.72	88.11	79.98	84.90	68.33	66.05	84.96
Gemini 2.5 Pro Preview (2025-03-25)	Google	76.69	87.53	71.08	89.16	62.47	69.31	80.59
Claude 3.7 Sonnet Thinking	Anthropic	74.50	76.17	73.19	79.00	69.11	68.27	81.25
o4-Mini Medium	OpenAI	74.40	78.47	74.22	81.02	68.47	62.41	81.83
Qwen 3 235B A22B	Alibaba	73.23	78.61	65.32	78.78	68.31	60.61	87.73
DeepSeek R1	DeepSeek	72.49	77.17	74.98	77.91	69.63	54.77	80.51
Qwen 3 32B	Alibaba	71.03	77.75	64.24	75.58	68.29	55.15	85.17
Grok 3 Mini Beta (High)	xAI	70.25	87.61	54.52	77.00	64.58	59.09	78.70
Gemini 2.5 Flash Preview	Google	69.93	73.47	60.33	81.80	65.53	59.43	79.02
QwQ 32B	Alibaba	69.50	76.72	61.36	76.08	69.53	51.48	81.83
GPT-4.5 Preview	OpenAI	65.93	54.42	76.07	67.94	60.07	64.76	72.33
Qwen 3 30B A3B	Alibaba	65.32	66.83	47.47	72.20	66.60	55.58	83.23
Claude 3.7 Sonnet	Anthropic	64.62	49.11	74.28	64.65	59.96	63.19	76.49
Grok 3 Beta	xAI	63.17	48.53	73.58	62.75	55.63	53.80	84.74
GPT-4.1	OpenAI	62.99	44.39	73.19	62.39	66.40	54.55	77.05
DeepSeek V3.1	DeepSeek	62.82	44.28	68.91	71.44	64.02	46.82	81.47
ChatGPT-4o	OpenAI	61.65	48.81	77.48	55.72	66.52	49.43	71.92
Gemini 2.0 Flash	Google	60.05	44.25	64.74	63.19	59.92	42.39	85.79

https://livebench.ai/#/ 1/2

Figure A.1: LiveBench results 13-5-2025 part 1

5/13/25, 11:20 AM LiveBench

Model	Organization	Global Average	Reasoning Average	Coding Average	Mathematics Average	Data Analysis Average	Language Average	IF Average
Qwen2.5 Max	Alibaba	60.03	38.53	66.79	56.87	64.27	58.37	75.35
GPT-4.1 Mini	OpenAI	59.05	53.78	72.11	58.78	61.34	38.00	70.31
Claude 3.5 Sonnet	Anthropic	57.94	43.22	73.90	50.54	56.19	54.48	69.30
LearnLM 2.0 Flash Experimental	Google	57.27	39.72	64.30	61.10	51.42	43.34	83.76
Phi-4 Reasoning Plus	Microsoft	56.64	57.83	60.59	62.83	54.74	30.69	73.17
Mistral Medium 3	Mistral AI	56.59	41.97	61.48	59.74	60.20	44.74	71.40
DeepSeek R1 Distill Llama 70B	DeepSeek	55.51	59.81	46.65	58.80	60.81	37.05	69.94
Llama 4 Maverick 17B 128E Instruct	Meta	55.19	43.83	54.19	60.58	47.11	49.65	75.75
Step 2 16K	StepFun	54.05	42.39	57.58	43.68	62.35	38.41	79.88
GPT-4o	OpenAI	53.95	39.75	69.29	41.48	63.53	44.68	64.94
Gemini 2.0 Flash Lite	Google	53.75	32.25	59.31	54.97	65.39	33.94	76.63
Hunyuan Turbos	Tencent	50.77	38.22	50.35	57.47	47.99	34.46	76.13
Mistral Large	Mistral AI	50.25	33.83	62.89	42.20	54.20	40.45	67.93
LearnLM 1.5 Pro Experimental	Google	49.30	34.86	58.93	56.71	39.30	37.86	68.16
Dracarys2 72B Instruct	AbacusAI	49.20	37.49	58.73	52.25	48.48	33.06	65.22
Qwen2.5 72B Instruct Turbo	Alibaba	49.04	34.08	57.26	51.88	50.16	36.47	64.39
Llama 3.3 70B Instruct Turbo	Meta	48.86	32.53	51.82	41.40	40.79	43.97	82.67
Gemma 3 27B	Google	48.44	34.42	48.94	52.27	38.80	41.31	74.90
DeepSeek R1 Distill Qwen 32B	DeepSeek	47.40	44.36	46.33	60.13	46.94	30.92	55.71
GPT-4.1 Nano	OpenAI	46.58	35.58	63.21	42.39	49.82	30.96	57.54
Dracarys2 Llama 3.1 70B Instruct	AbacusAI	46.47	36.67	41.14	40.30	55.13	42.37	63.24
Mistral Small	Mistral AI	45.92	37.08	49.65	38.39	52.14	34.59	63.66
Claude 3.5 Haiku	Anthropic	44.98	26.19	53.17	34.84	54.12	39.71	61.88
Nova Pro	Amazon	44.33	28.25	49.65	37.70	44.34	38.94	67.13
GPT-4o Mini	OpenAI	43.41	25.64	55.02	38.05	55.10	29.88	56.80
Nova Lite	Amazon	39.11	32.00	45.04	34.62	41.24	27.62	54.13
Command R Plus	Cohere	34.88	21.64	27.13	22.82	49.23	30.86	57.61
Qwen2.5 7B Instruct Turbo	Alibaba	34.37	22.31	34.29	36.81	42.33	18.38	52.11
Nova Micro	Amazon	33.67	25.42	28.92	34.15	41.29	24.19	48.04
Command R	Cohere	31.39	20.58	26.10	18.35	39.77	27.93	55.62

BibTeX

```
@inproceedings{livebench,
  title={LiveBench: A Challenging, Contamination-Free {LLM} Benchmark},
  author={Colin White and Samuel Dooley and Manley Roberts and Arka Pal and Benjamin Feuer and Siddhartha Jain and Ravid Shwartz-Ziv and Neel Jain and Khalid Saifullah and S
  booktitle={The Thirteenth International Conference on Learning Representations},
  year={2025},
}
```

Colin White*[1],Samuel Dooley*[1],Manley Roberts*[1],Arka Pal*[1],
Ben Feuer[2],Siddhartha Jain[3],Ravid Shwartz-Ziv[2],Neel Jain[4],Khalid Saifullah[4],Siddartha Naidu[1],
Chinmay Hegde[2],Yann LeCun[2],Tom Goldstein[4],Willie Neiswanger[5],Micah Goldblum[2]
[1]Abacus.AI,[2]NYU,[3]Nvidia,[4]UMD,[5]USC

The site was inspired by the Nerfies project and by LiveCodeBench.

https://livebench.ai/#/ 2/2

Figure A.2: LiveBench results 13-5-2025 part 2

Appendix B

University of Amsterdam LLM Policies

https://www.uva.nl/en/about-the-uva/about-the-university/ai/ai-policy/ai-policy.html

UvA policy on AI

Developments in the field of AI move extremely fast, which is why the UvA regularly reviews its AI policy. We are interested in the potential of AI programs in the fields of education and research. In the development of our AI policy, the current focus is on the role of AI in education.

AI in reseach

In developing our policy, the focus is currently on AI in education. When a policy for AI in research has been developed, we will post it on this page. Are you a UvA researcher? Please check with your research institute to see if any policies apply to AI research.

AI in education

The UvA's AI policy for education (in Dutch) includes action plans for the short, medium and long term.

Short term: Supporting lecturers and knowledge sharing

While some lecturers are enthusiastic about the potential opportunities offered by AI, for others AI is a cause for concern. For a large group of lecturers, the sudden emergence of ChatGPT has created extra work, mainly because they have had to adjust their assessments to prevent fraud. Accordingly, in the short term, the UvA policy should focus on supporting degree programmes in general and lecturers in particular in dealing with AI in their teaching and assessment.

Tips for lecturers on dealing with AI in education

Medium term: Future-proof education

For all education offered by the UvA, it will be necessary to assess whether programmes' exit qualifications are still appropriate for this new reality. Do the current exit qualifications still allow us to prepare people for the current and future labour market? Or do they need to be refined? An important question that will play a role here is: what are the skills we want students to have when they leave university?

Long term – Anticipating the future

Although the precise outlook is still uncertain, it is clear that the impact of AI on education will be huge and ongoing. To gain better insight, we've partnered with VU Amsterdam to establish an AI task force.

Learn more about the task force

https://student.uva.nl/onderwerpen/ai-tools-en-je-studie

AI-tools en je studie

Laatst gewijzigd op 19-10-2024 16:10

print

Lees wanneer je generatieve AI-tools zoals ChatGPT wel en niet mag gebruiken voor je studie en wat de mogelijke risico's zijn.

list-ulInhoudsopgave

Minimaliseer

Studenten en docenten aan de UvA mogen AI-tools zoals ChatGPT (nog) niet actief inzetten voor onderwijs en toetsing. Wel verkent de UvA het gebruik van artificiële intelligentie in het onderwijs in de toekomst.

Over ChatGPT

ChatGPT is een taalmodel, getraind om snel tekst te genereren en vragen te beantwoorden gebaseerd op de prompts die jij als gebruiker invoert. Als gebruiker kun je reageren met vervolgvragen of -opdrachten, om de output verder af te stemmen op jouw wensen.

Waarom wordt ChatGPT niet ingezet bij het onderwijs aan de UvA?

ChatGPT is een product van een commercieel bedrijf. Op dit moment is onduidelijk wat er met gegevens van gebruikers en de ingevoerde data wordt gedaan. Dat maakt de tool (nog) niet geschikt voor gebruik binnen de UvA. Je docent kan je daarom niet verplichten om ChatGPT te gebruiken bij studieopdrachten.

Gebruik AI-tools niet om teksten te schrijven die je gaat inleveren

Als UvA-student mag je kwalitatief hoogwaardig en innovatief onderwijs verwachten. Daar hoort een intrinsiek gemotiveerde houding bij. Het blijft belangrijk om zelf schrijfopdrachten te maken, en dit niet door een generatieve AI-tool te laten doen. Zo train je academische vaardigheden, die onmisbaar blijven tijdens je verdere opleiding en op de arbeidsmarkt.

Niet-zelfgeschreven werk kan worden aangemerkt als fraude

In de UvA fraude- en plagiaatregeling staat vermeld dat docenten de kennis, het inzicht en de vaardigheden van een student moeten kunnen beoordelen. Van studenten verwachten we dan ook dat zij volledig zelfgeschreven werk inleveren. Tenzij nadrukkelijk anders vermeld, is gebruik van AI-programma's zoals ChatGPT daarom niet toegestaan. Het inleveren van werk dat je niet zelf hebt geschreven, kan daarom worden aangemerkt als fraude. In het geval van fraude treedt de UvA streng op.

Risico's bij het gebruik van AI-tools

AI-tools kunnen je helpen om effectiever te studeren. Zo kun je ChatGPT gebruiken voor een brainstorm, kan ChatGPT je overhoren voor je tentamen of een tekst voor je vertalen. Het kan een handige assistent zijn die altijd beschikbaar is. Maar, het is belangrijk om je bewust te zijn van de risico's.

- Antwoorden kunnen plausibel klinken en toch onjuistheden bevatten. Zo genereert ChatGPT output op basis van kansen en statistiek. ChatGPT controleert niet of de informatie feitelijk juist is.
- De output kan schadelijke vooroordelen en stereotypen bevatten. De datasets die ChatGPT gebruikt, zijn bijvoorbeeld niet representatief.
- Veel AI-tools slaan jouw interacties op. Dat brengt risico's met zich mee op het gebied van privacy en kennisveiligheid. Voer daarom geen

privacygevoelige of andere vertrouwelijke informatie in, zoals vertrouwelijke onderzoeksgegevens, patiëntgegevens of persoonsgegevens van medestudenten of docenten. Ook een betaald account biedt geen garanties voor veiligheid en privacy.

- Maak geen gebruik van je UvA-account. Gebruik een dummy-mailadres dat niet tot de UvA en liefst ook niet tot jou als persoon herleidbaar is.

E-module voor studenten

Het Teaching & Learning Centre ScienceExterne link heeft, in samenwerking met dr. Jelle ZuidemaExterne link, een Engelstalige e-module voor studenten ontwikkeld over het verantwoord omgaan met AI-tools (en met name ChatGPT) in het hoger onderwijs. Deze interactieve e-module bestaat uit teksten, kennisclips en korte kennisquizzen. Het duurt ongeveer 45-60 minuten om de e-module te doorlopen.
Ga naar de E-module

https://student.uva.nl/onderwerpen/plagiaat-en-fraude

Plagiaat en fraude

Laatst gewijzigd op 19-10-2024 16:12

print

De UvA treedt streng op tegen plagiaat. Wie ernstige fraude pleegt, kan worden uitgeschreven en de studie niet meer afmaken. Zorg er dus voor dat je de academische regels kent.

list-ulInhoudsopgave

Minimaliseer

Plagiaatcontrole in Canvas

Docenten kunnen in Canvas plagiaatcontrole instellen voor opdrachten. Het werk dat je hier inlevert wordt met de software Turnitin Similarity gecontroleerd op plagiaat. Het programma vergelijkt jouw tekst met bronnen op het internet, tijdschriften en (eerder) ingeleverd werk van mede UvA-studenten. Hieruit komt een percentage voort dat aanduidt in hoeverre jouw ingeleverde opdracht overeenkomt met bestaand werk. De docent ontvangt een overzicht van de resultaten en krijgt hiermee een indruk van de originaliteit van je werk.

Een opdracht met plagiaatcontrole inleveren

1. Open je opdracht in Canvas
2. Als de docent plagiaatcontrole heeft ingesteld, word je gevraagd om na het toevoegen van je tekst te bevestigen dat de opdracht jouw eigen, originele werk is.
3. Vervolgens kun je de opdracht inleveren

Een plagiaatrapport openen en interpreteren

De docent kan instellen dat studenten ook toegang krijgen tot hun plagiaatrapport. Zij doen dit alleen als het toegevoegde waarde heeft voor het onderwijs en zullen je hierover informeren. Op de website van Turnitin Similarity vind je o.a. instructies voor het openen van het plagiaatrapportExterne link en het interpreteren van de overeenkomstscore.Externe link

Bescherming persoonsgegevens

Het intellectueel eigendom van je werk is gewaarborgd. De bestanden en bijbehorende gegevens worden alleen bewaard om plagiaatcontroles uit te voeren en eventuele gevallen van plagiaat te onderzoeken. In geen geval wordt Turnitin eigenaar van documenten. Het bedrijf moet zich contractueel houden aan de Algemene Verordening Gegevensbescherming (AVG)

Fraude en Plagiaat Regeling Studenten UvA

Raadpleeg voordat je een tentamen, paper of essay maakt de Fraude en Plagiaat Regeling Studenten UvAExterne link. Hierin staat precies omschreven wat de universiteit onder plagiaat en fraude verstaat, en aan welke regels je je moet houden.

Auteursrecht

Als student maak je vaak gebruik van materiaal van derden. Maar mag je zomaar citeren uit andermans werk, of een op internet gevonden afbeelding gebruiken voor je scriptie? De Auteurswet staat toe dat gedeelten van een werk gebruikt worden zonder dat er toestemming nodig is van de auteursrechthebbende, maar wel volgens bepaalde regels. Meer informatie vind je op de pagina van de bibliotheekExterne link.

Kijk ook bij

- Onderwijs- en examenregelingen (OER) en andere regelingen
- Examencommissie
-

https://www.uva.nl/en/about-the-uva/about-the-university/ai/ai-in-education/vu-uva-task-force-on-ai-in-education/vu-uva-task-force-on-ai-in-education.html

VU-UvA task force on AI in education

The task force advises the Executive Boards of both institutions on the use of AI in education. It does this in accordance with the principles of the Netherlands Code of Conduct for Research Integrity. The task force advises on the challenges, as well as the new opportunities that artificial intelligence offers education.

What does the task for do?

The task force:

- Highlights recent developments;
- Advises on the adjustment and tightening up of policies;
- Provides a set of criteria that software must meet in order to ensure academic integrity.

The task force has produced three recommendations on the use of AI in education.

Task force principles

- **Transparent to the outside world**
- **Transparent policies for the organisation**
- **Transparent about expectations**

- **Transparency of software vendors**

Composition of the task force

Participants from the UvA include:

- Prof. Peter van Baalen, expertise: innovation models
- Prof. Natali Helberger expertise: Artificial Intelligence
- Dr. Marjolein Lanzing expertise: Philosophy of technology
- Dr. Jelle Zuidema, expertise: Natural Language Processing

Participants from the VU include:

- Dr. Peter Bloem, expertise: Artificial Intelligence
- Dr. Ilja Cornelisz, expertise: Education Sciences
- Prof. Felienne Hermans, expertise: Computer Science
- Prof. Piek Vossen, expertise: Computational Lexicology