OPEN ACCESS

GRS Journal of Multidisciplinary Research and Studies

Abbreviate Tittle- GRS J Mul Res Stud

ISSN *(Online) - 3049-0561*

https://grspublisher.com/journal-details/GRSJMRS

Vol-3, Iss-5 (May)-2026)

Generative AI and the Illusion of Learning: Examining Metacognitive Fluency and Knowledge Overestimation in Large Language Model-Assisted Education

Abdul Alim*

Department of Business Development, National Research University of Higher Economics (HSE).

Article History

Received: 10.04.2026

Accepted: 05.05.2026

Published: 30.05.2026

Corresponding Author:
Abdul Alim

Abstract: Fast spread of smart computer programs, especially big language models, changed how people learn today - offering quick answers, short recaps, because they help solve problems fast. Even though such tools bring clear advantages in education, there's a risk learners might feel they know more than they actually do. What looks like solid grasp could just be surface-level comfort. This work digs into how thinking about one's own thinking connects to overconfidence when using AI helpers. Ideas around mental effort, memory recall, reliance on outside aids play a role here. Smoothness of machine-made replies often tricks students into believing they've learned deeply. That smooth flow gets mistaken for real insight, which leads them to rate their knowledge higher - even if test results tell another story. Using large language models too much might make people think less deeply, skip useful memory exercises, then forget things faster. What students believe they learn often does not match what they actually retain - a pattern showing how easily judgment can drift off track with AI help. Even so, artificial intelligence tools still hold value inside classrooms if paired with methods like questioning oneself or pausing to reflect. Clear guidance emerges near the end: those who study, teach, or build tech should adjust their habits to support sharper thinking alongside machines. This work adds weight to ongoing conversations about pairing smart devices with mental effort, aiming for deeper understanding that lasts longer.

***Keywords**: Generative Artificial Intelligence (AI); Large Language Models (LLMs); Illusion of Learning; Metacognitive Fluency; Knowledge Overestimation; Judgment of Learning (JOL); AI-Assisted Learning; Cognitive Offloading; Retrieval Practice; Educational Technology; Metacognition; Learning Outcomes.*

Cite this Article

A. Alim, (2026) Generative AI and the Illusion of Learning: Examining Metacognitive Fluency and Knowledge Overestimation in Large Language Model-Assisted Education *GRS Journal of Multidisciplinary Research and Studies, Vol-3 (Iss-5).22-35*

Introduction

1.1. Background of Study

Out of nowhere, AI that writes and thinks like a person has changed classrooms in deep ways. Not long ago, big language machines - ones like - started showing up in student routines, doing more than just helping here and there. Instead of only searching facts, they now explain tough ideas on the spot. Sometimes within seconds, answers appear clear even when topics feel tangled. Because these tools respond fast and make sense, many find them hard to ignore during homework time. Slowly but surely, getting help from smart software became normal while studying. One day at a time, using machine-made summaries or step-by-step solutions shifted from rare to routine. Knowledge no longer comes only from books or teachers - it flows through digital helpers too. What once took hours to grasp can now unfold quickly with a few typed lines. Behind every search, behind each prompt, learning quietly evolves without announcements or warnings.

Even though large language models can help education in big ways, more people using them brings up real questions about how well students actually learn and whether they truly know what they understand. Studies in how people learn have long found that folks often misjudge just how much they get something. It is common for learners to feel confident when they shouldn't - mistaking familiarity or clear wording for true grasp. Smooth reading tricks minds into thinking knowledge is solid. That smoothness, called meta-cognitive fluency, gives a warm impression of knowing - but warmth isn't proof. Easy processing might feel like mastery, yet rarely means ideas stick or make sense later.

Looking at regular classrooms, research shows just watching solved problems boosts confidence but leads to shallower learning when besides tackling questions yourself. Because answers from generative AI come fast, sound smooth, feel tailored, and chat like a person, the gap might grow. That ease can trick someone into thinking they grasp an idea fully, even if they do not. So belief in one's knowledge climbs, yet real mental effort drops. Mastery seems within reach - until faced with new situations where old shortcuts fail.

Even though more researchers are looking into generative AI in schools, much of the work zeroes in on cheating, speed, access, or grades. Yet when it comes to how using AI affects thinking habits or self-awareness in learning, insights remain thin. One idea - that answers made by large language models might create a false sense of understanding - hasn’t been tested directly through controlled study.

Looking into those who use AI-made answers, do they think they get it better than those working things out themselves? What ties exist between using smart machines, how people judge their own grasp, and what they really learn? Through studying these new mental patterns, the work adds weight to talks about school settings shaped by machine thinking. When tools shape thought, questions arise - whose knowledge counts, and where does real knowing begin? These moments of confusion may point toward deeper shifts in how minds grow alongside circuits.

1.2. Significance of The study

Fast changes in how schools use smart computer tools have opened fresh paths for studying, getting help with schoolwork, and reaching information more easily. Tools like big talking machines - able to write essays or explain tough ideas - are now common among learners who want answers fast. Because these systems respond quickly, some rely on them heavily when working through assignments. One quiet risk? Believing you know something just because a machine showed it clearly. That feeling may not match real skill. What looks like mastery could be mistaken confidence instead. Research often highlights benefits but skips deeper mental effects hiding beneath regular use. Few studies check how trust in AI might distort self-awareness about what someone truly grasps. This work matters since it tackles that hidden gap - examining if ready-made replies trick minds into thinking they’ve learned more than they actually have.

One step at a time, this work adds weight to educational psychology through fresh insights about thinking awareness and how smoothly information feels when handled. Earlier findings showed people tend to mistake mental comfort for real mastery. Yet those results mostly came from standard classroom settings - like studying solved problems or going over class notes. Now, placing these ideas into the evolving space of AI-driven learning shifts things forward. Through controlled testing, it separates what happens when answers come from large language models versus confidence in one's own grasp. That gap reveals something new about biased thinking patterns as they play out while using smart tools online.

Another reason this work matters: it connects closely to the fast-moving world of artificial intelligence in schools. Right now, talk about generative AI centers on cheating, copied work, speed, and tools working better. Important as those topics are, they tend to miss what happens inside a student’s mind when figuring out how well they’re learning. Instead of grades or test scores, this project looks at mental patterns behind using AI while studying. Figuring out if smooth answers from machines lead to misplaced certainty becomes key - believing too much in one’s understanding might weaken effort, deeper analysis, or solving problems without help.

One key takeaway stands out - confidence boosted by AI answers might backfire. When students feel too sure after seeing machine-made work, real practice often slips away. Instead of digging deep, minds switch off, relying on quick fixes. That shift chips away at lasting memory, problem-solving reach, and clear thinking over time. Spotting this risk early opens space for smarter teaching moves. Lessons could lean into doing, not just watching, pulling students into the act of figuring things out. Schools gain a chance here - to shape rules that keep AI in its place as a helper, never a substitute. Real grasp matters more than surface-level nods. Letting tools guide too much dulls the mind's edge. Learning stays alive when effort remains central.

What makes this study matter? It might shape how we build smarter tools for learning down the line. Should findings show that machine-made explanations trick learners into false confidence, creators can adjust their designs - baking in moments for reflection, self-checks, or hands-on problem solving instead. Picture a system that holds back answers just long enough for students to try on their own first. Tweak by tweak, these changes may help curb overestimation while nurturing deeper understanding.

Right now, more people are using generative AI when making choices at work, school, or home. Still, there's a risk it might give learners false confidence about their real abilities. Later on, those same students might move into jobs thinking they know more than they really do. That gap - between what someone thinks they understand and what they truly grasp - can quietly influence how well they perform. It may even shape the way decisions get made over time. Learning doesn’t stop after school ends, so flawed foundations matter beyond classrooms. Because of this, looking closely at how minds change when leaning on AI isn’t just useful for schools. The ripple goes further. Everyone stands to feel its effect.

What makes this work stand out lies in how it approaches uncharted territory. While some experts suspect generative AI can lead people to feel they understand more than they do, hardly any tests have pinned down that effect clearly - especially when comparing those who rely on AI tools with those who form ideas on their own. It’s exactly this missing piece the project tackles head-on. Fresh data comes into play here, adding something rare to ongoing conversations about AI's role in thinking and learning. Outcomes could quietly open doors for later work weaving together insights from education theory, mental processing models, digital interface design, and machine reasoning studies.

This research matters simply because it looks into something new - how people learn when tech changes fast. Since artificial intelligence now alters classrooms everywhere, thinking about its impact on thought processes feels more urgent than before. What happens inside a learner's mind today can shape how knowledge sticks tomorrow.

1.3. Problem Statement

Fast spread of AI that creates content is reshaping how pupils find, handle, then interact with school materials. Tools like ChatGPT, deep seek now regularly produce detailed answers, breakdowns, besides condensed versions in many topics. Though they make studying easier and faster, what these systems do to thinking patterns or personal judgment still lacks clear insight. One big idea from studying how people learn is that students sometimes think they get something when they really do not. Earlier work shows folks tend to misjudge what they know, especially if facts feel smooth or come in neat packages. When ideas flow easily, it tricks the mind into feeling confident even without real mastery. That false sense of knowing - called metacognitive fluency - can make someone sure about knowledge they lack. Yet nearly all past experiments looked at old-school settings like textbooks or step-by-step lessons handed out by teachers.

Out of nowhere, generative AI shows up changing how thoughts flow. Solutions made by machines often look clean, fit together neatly, appear fast - this smoothness might lighten mental load when picking up something new. Because it feels so easy to follow what comes from AI, people could believe they grasp ideas better than they really do. Here lies a real puzzle for study: does seeing answers built by large language models push learners into feeling too sure about their knowledge? Especially when stacked against those who work things out on their own.

Even though more people are looking into how generative AI can be used in education, much of today's research centers on grades, speed of work, along with worries about cheating or improper use. Little has been written about the mental processes involved when students learn with help from AI - especially how content made by large language models affects their ability to judge their own understanding. Without deeper insight, it's hard to say if using AI while learning leads to a new kind of thinking error - one that older ideas about learning fail to capture. So here's the core issue this research tackles: we do not have solid proof showing if using generative AI tricks students into feeling more confident about what they know than they should be. The focus zeroes in on comparing those who lean on answers created by large language models against others working out problems themselves - watching closely how sure each group feels versus what they actually remember and apply later. Getting clear on this matters because it shapes how educators think about the hidden mental costs when placing AI tools into classrooms.

1.4. Aims and Objectives of the Study

One goal here stands out: checking if using AI-made answers - like ones from ChatGPT - affects how students judge what they know. When people see responses built by big language models, something shifts inside their thinking. Instead of working things out alone, some rely on ready-made replies. This shift might trick minds into feeling more confident than they should be. Solving problems solo usually brings clearer awareness of gaps. But reading smart-sounding outputs can create false clarity. The ease of grasping someone else's answer feels like real learning. That smooth mental ride may distort self-assessment sharply. Confidence swells without actual mastery growing. Working through steps personally builds truer insight. Relying on generated text skips struggle - and skips growth too.

What happens inside the mind when people learn with help from artificial intelligence? This research looks closely at thinking patterns tied to that process. Ease of grasping material plays a big role, shaping how students assess their own knowledge. Confidence does not always match real results. The way tasks feel simple can shift both awareness and outcomes. Learning gains depend heavily on these mental shortcuts. Feelings of clarity may lead some to overestimate mastery. Actual progress sometimes lags behind self-judgment. Perception acts as a filter between effort and achievement.

1.5. Objectives of the Study

For this goal, the work follows these clear targets

Looking into how answers made by large language models shape what people think they understand

Looking at how students feel after working through AI-made answers, the research checks if they believe they understand more or feel surer than peers tackling tasks alone. Instead of guessing outcomes, it measures self-reported clarity and belief in their grasp when support is given versus none at all. Some tackle questions using machine help, others do not - feelings on learning are then compared across both paths.

To compare actual learning performance between AI-assisted and self-generated learning groups

One way to look at it is through how students handle problems right after seeing answers made by large language models. Does their work get better, worse, or stay about the same? The research aims to find that out by watching real attempts unfold. What happens when they face new questions soon afterward becomes part of the picture. Performance shifts might show up - then again, maybe not. Close observation tells more than assumptions ever could.

Checking how well information sticks later on

One goal of the study is checking if students using AI answers later remember less than those working out answers on their own. Instead of creating responses, some simply take what AI gives them - how that affects memory shows up days afterward. While one group solves problems independently, another leans entirely on machine help without adding much thought. Memory strength weeks later becomes a clue about learning depth. Because thinking through steps changes how knowledge sticks, skipping mental effort might weaken long-term results.

To measure transfer of learning to novel problems

One way to start is by looking at how people handle fresh problems after seeing answers made by artificial intelligence. Some folks get ideas from machines. Others work things out on their own. The research checks if those who view machine-made responses can still use what they picked up when faced with situations they have never seen before. A different path unfolds when comparing these two groups. Outcomes may shift depending on where the original insight came from. New settings test whether borrowed methods stick or fade.

To investigate the relationship between meta-cognitive fluency and confidence

Looking at how clear AI answers feel might skew how students judge their own understanding. Smooth explanations can trick the mind into feeling more confident than it should be. When things sound simple, people often think they get it - maybe too quickly. This project checks if that ease creates a false sense of knowing. What seems fluent may not always reflect real learning underneath.

To identify potential overconfidence bias in AI-assisted learning

Looking into how students who rely on answers made by large language models might misjudge their own skills compared to how they truly perform forms the core of this study.

To contribute empirical evidence to AI-in-education literature

Ending things off, researchers want hard data from tests about how generative AI affects students' thinking patterns in school environments - how they understand their own learning, see knowledge growth, while also tracking habits like skewed judgment that might quietly take root.

1.6. Research Questions

The questions are about looking into how meeting generative AI - like answers built by big language models - might create a false sense of knowing. When students see solutions made by tools like ChatGPT, their minds may feel fluent, mistaking smoothness for real grasp. This ease could twist both self-judged comprehension and true knowledge gained. Because of these concerns, along with the core idea and guiding theory, certain questions start forming. Each one digs into effects shaped by interaction with machine-made explanations

Main Research Question

Could we find answers made by large language models trick to someone into feeling they understand more than they do, especially when contrasted with working out problems on their own?

Sub-Research Questions

1. What happens to a student's sense of grasp when they review AI-made answers instead of working out their own?
2. What happens when people learn by working out answers themselves instead of seeing ready-made ones from an AI?
3. One group gets solutions handed to them. The other tries to figure things out alone. Does one way lead to better understanding than the other?
4. When students rely on AI answers, does their self-assessment match how well they truly perform?
5. What happens to memory when people use generative AI instead of creating ideas on their own?
6. Over weeks, does one method stick better than the other?
7. Some rely on AI tools while others think through things alone - how do those paths differ later?
8. When learning sticks around, is it because of how it was built? One path uses machines, the other builds from within.
9. How much does ease of understanding shape how sure students feel after seeing answers made by large language models?
10. Could working with answers made by large language models make it harder for students to apply what they know when facing new kinds of problems?

1.7. Research Gap

Even though more people study artificial intelligence in schools, few look closely at how using tools like Chat GPT, Cloud AI, Deepseek affect thinking about one's own learning. These systems often help students explain ideas or work through homework questions. Yet instead of asking what happens inside a student's mind, researchers usually just check test scores afterward. Behind each correct answer might be confusion masked by confidence - something rarely measured. What changes when a learner trusts an AI too much becomes harder to track without studying awareness alongside results. Tools shape thought quietly, sometimes altering judgment without notice. Understanding that shift means watching both belief and ability together - not only whether answers improve but why they feel right.

Much research in education science looks at self-aware thinking, how smoothly people handle new ideas, and false confidence during standard classroom learning. When facts seem clear or students see step-by-step solutions, they tend to believe they know more than they do. Yet most of what we know comes from settings using printed texts, spoken lessons, or handouts made by teachers - not material created by artificial intelligence. Still, one fact remains: the mind tricks itself just as easily online. Meanwhile, work in AI for education mostly focuses on things like keeping exams fair, stopping dishonesty, boosting how fast students learn, holding attention, along with creating custom teaching software. Even though such efforts show artificial intelligence might make knowledge easier to reach and assist schooling, they rarely explore if these smart tools shift a learner's inner thinking habits - especially skills in judging just how well they truly understand something.

It turns out that hardly any studies have tested how using AI help lines up against working things out alone when it comes to judging your own understanding. Instead of building answers on their own, some students just take what smart machines give them - yet we do not know if this leads to thinking they get it better than they really do. That false sense of grasp caused by slick AI replies has been talked about in theory, though nobody has cleanly measured it under proper test conditions so far.Beyond that, few past investigations look closely at how feeling informed compares to truly grasping material when using AI help. While many projects track results like exam marks or finishing assignments, they often skip checking if a person's belief in what they know matches what they actually retain. That leaves an open question - could generative AI quietly promote shallow understanding, making things feel clear without deep mental effort?

Long-term impacts remain poorly studied - like how well people keep or apply what they learn. Even if artificial intelligence helps users complete tasks faster today, we do not know if those gains stick around later or simply vanish like a quick fix. What also stays hidden is how smoothly processed answers from machines mix with personal study habits; sometimes ease might weaken the drive to push through hard thinking on your own. So here's what's missing so far - real proof that generative AI tricks people into thinking they understand more than they do. That false sense could inflate confidence while quietly lowering precision in judging one's own grasp. To tackle this, the work tests how encountering answers made by large language models shapes someone's belief about their comprehension. It also checks true test results, memory after time passes, and ability to apply ideas later. All of these get weighed against experiences where learners build answers themselves instead.

1.8. Theoretical Framework

One reason this research exists lies in how people think about thinking - especially when they guess if they know something or just feel like they do. Sometimes it feels easier to process ideas, which tricks students into believing they understand more than they actually do. Built around that idea are three lines of thought: one explores inner awareness during learning, another looks at mental ease, while a third follows how individuals manage their own progress. Together, these help make sense of what happens when tools like advanced AI chat systems step into education. When learners rely on smart machines that generate answers, those internal judgments shift - quietly but clearly.

Meta-cognition Theory

Thinking about helping people notice how they learn. Some call it knowing what you know; others see it as tracking your efforts while studying. One moment you might feel certain, yet later realize confusion set in earlier. This ongoing check isn't flawless - mistakes sneak through gaps in judgment. Still, noticing errors matters more than pretending perfection exists. Surprisingly, thinking you get it just because an answer seems clear can trick your mind during AI-supported study sessions. Relying on neatly built responses from large language models might make a student feel confident - yet real grasp often doesn't come without personal mental work. A sudden sense of knowing isn't always actual knowledge, especially when help comes pre-packed and polished.

Processing Fluency Theory

Smooth thinking shapes what feels true. When words slide into the mind without effort, they often seem truer - even if nothing has really been grasped. Clarity tricks judgment, mistaking speed for insight. Simple layout, clean fonts, familiar ideas - these slip through thought like water. The brain confuses fast processing with solid knowledge. Effortless does not mean correct, yet it plays the part well.

Smooth talk from large language models can feel convincing. Because it flows so easily, people might think they get something just because it sounds clear. When explanations slide into the mind without effort, confusion gets mistaken for clarity. Ideas seem solid when they're really just neatly arranged. The brain likes things that fit together fast - too fast sometimes. What reads effortlessly might leave gaps behind. Easy words do not always mean deep learning has happened.

Self-Regulated Learning Theory

Starting strong, some students manage their own study by deciding what to achieve, checking how they're doing, then judging results afterward. These learners dive into tough thinking tasks, quiz themselves now and again, plus pause to think about what stuck. It breaks down though - whenever outside aids take over, mental work fades fast.Most times, students lean too hard on AI answers when help is around. Because of that, they skip wrestling with tough ideas themselves. Without that push, checking their own understanding gets harder. Overconfidence sneaks in, even if knowledge isn't solid. Memory fades faster down the line. Mastery feels real, but it rests on shaky ground. Real learning needs effort - they miss that part.

Framework of the Study

Putting these three ideas together, the research suggests this:

- ✓ Smooth answers from large language models slide into your mind without effort. Because they flow so well, what you read seems clear right away.
- ✓ Folks often think they get something better than they do when it feels smooth to process. That ease tricks them into trusting their grasp too much.
- ✓ When thinking less, people struggle to guide their own learning well. This leads to lower involvement in tasks. Memory suffers as a result. Focus fades without steady mental effort.

Some mix of these processes hints that generative AI could produce a special kind of mental trick. Learners might mistake how clear something feels for real mastery. Because of this blend, the model sets up a way to study if seeing AI-made answers skews both belief in knowing and true performance. The whole setup helps probe whether repeated contact warps self-assessment along with results.

1.9. Limitation of This Study

Even if this work tries to carefully examine how using generative AI affects thinking about one's own learning, it still has some weak spots. Some gaps come from how narrow the setup was, tools that couldn't capture everything, differences among people who took part, along with how fast things like big-word software - including programs similar to DeepSeek - keep shifting.

1. Limited Generalizability

One reason these results might not apply widely is the type of people included. Suppose most participants are college students - then what works here could miss how teens learn elsewhere. Picture someone trained on the job; their experience with learning tech differs sharply. Think about it: knowing how to think about thinking changes when culture shifts. Even comfort with artificial intelligence tools isn't the same everywhere you look.

2. Artificial Experimental Environment

When research happens in tightly managed labs, classroom reality often slips through the cracks. Real learners mix textbooks, videos, chats with classmates, and dip into AI help across days. Tight setups miss these layers, flattening messy but true patterns of how people actually learn. That gap can weaken how well findings fit outside lab walls.

3. Measuring Learning Outcomes over a Short Period

Right away, some lab tests check how well people perform, feel they get it, or remember right after. Yet real mastery usually shows up later, when facts stick or apply elsewhere weeks on. When follow-up quizzes happen too soon or skip longer waits, those experiments miss deeper shifts from using AI in studying.

4. Rapid Advances in Generative AI

Fresh waves of change keep reshaping how well generative AI can understand, decide, and respond. Because these tools advance so quickly, today's results might not hold up when better versions arrive. Shifts in how clearly AI explains itself, tailors responses, or supports learning could reshape when and how meta-cognitive ease plays out. Later models may behave differently simply because they work differently under the hood.

5. Measuring How People Think About Their Thinking

Now think about how people judge their own grasp of a topic - often it comes down to asking them directly, though that route isn't always reliable. Sometimes someone rates their insight too high, sometimes too low, not because they learned more or less but due to quirks like eagerness, mood, or past experience rating themselves. Because of these personal filters, numbers drawn from such reflections can skew how we interpret awareness of one's thinking.

6. Prior Knowledge and Individual Differences

Some learners already know more, think differently, or have solved similar problems before - this shapes how well they do and how much they believe they grasp. Though people might be split into groups at random, personal traits can alter their reactions to answers made by AI, muddying outcomes. What one person sees as clear another finds confusing. Background matters even under controlled conditions. Past exposure shifts perspective quietly.

Unexpected mental shortcuts appear when least expected. Familiarity skews perception without notice.

Literature Review

Lately, attention's shifting. More researchers focus on AI's role in reshaping school habits - how kids think, act, respond. Tools such as Chat GPT? They've altered answer-seeking routines. That shift comes from new ways learners engage with facts. Despite shelves filled with writings on tech-assisted lessons, deep dives into self-awareness during problem-solving stay rare. Past findings pull threads from four angles. Judgment about personal understanding opens one door. Ease of grasping material links to flawed logic - that path matters too. Stepwise demonstrations help minds follow better. Then there's fresh evidence: generative models inching into daily classroom life right now.

1. Thinking About Thinking and Judging Learning

Surprisingly clear moments often hide shaky understanding beneath. People believe they know, though evidence suggests otherwise. Confidence tends to stretch further than accuracy ever reaches. This mismatch has earned a label all its own surprisingly, people often think they know more than they do. When tasks feel smooth or content seems simple, mistakes creep in without warning. Research into how students learn reveals a pattern - confidence grows too fast under slick conditions. Ease tricks the mind into false security now and then. Most people have heard of the Dunning-Kruger effect - where those less skilled in an area often think they're better than they really are. Still, even capable students can misjudge how much they know if what they study feels too neat or oversimplified. When it comes to judging their own learning, studies show folks lean on things like how easy a text seems or how familiar certain ideas feel. Instead of digging into actual comprehension, they take shortcuts based on first impressions.

Surprisingly, people often trust how clear something sounds rather than whether they actually understand it. When explanations feel smooth and neat, the mind can trick itself into feeling confident. Without realizing it, someone might think they've learned deeply just because words fit together well. This illusion grows stronger when thinking feels effortless. What seems like insight could simply be polished phrasing. Confidence builds not from knowledge but from surface traits. The brain prefers fluency over truth. A slick explanation doesn't mean real learning happened.

2. How Easy Thinking Shapes Hidden Biases

Easy thinking shapes opinions more than we might guess. The smoother something feels to understand, the truer it often seems. People tend to trust what slides effortlessly into the mind. Familiarity grows when mental work fades into background. Simple handling of ideas links closely to belief. Studies show this pattern again and again - whether recalling facts, weighing choices, or drawing conclusions. Smooth reading tricks the mind into trusting ideas more. Take this: when words flow fast, people believe them more - truth does not matter. School stuff that feels easy boosts how sure students feel. But remembering it later? Or using it somewhere new? That part often stays unchanged.

Right now, generative AI stands out because of how naturally it speaks. Large language models craft answers meant to flow - clear grammar, tight logic, smooth connections between ideas. That smoothness helps people engage - but there's a snag: ease can trick the mind into believing knowledge is deeper than it is. So when students see neat replies from AI, they might think they get it. Confidence grows fast - even if real grasp lags behind.

3. Worked examples and passive learning effects

Looking at how students learn from solved problems reveals a lot about their thinking. Step-by-step solutions, often seen in math and science classes, show exactly how answers are built. These guides tend to help beginners grasp ideas faster than trial-and-error methods. Evidence across many studies points to clearer understanding when learners study completed examples instead of solving everything themselves. Still, the benefit of worked examples comes with downsides. Even if people seem to grasp things at first, they usually find it hard to use that understanding in new situations later on. One reason sits in how just watching steps unfold keeps minds less engaged, leaving little room for real thinking work.

Surprisingly, some students feel they've got it down just by watching solved problems - yet without explaining things aloud or testing themselves, that confidence can be misleading. Oddly enough, what feels like smooth progress might actually hide shallow grasp, tying back to how poorly we sometimes judge our own learning. The brain tricks us: when something seems easy, we assume it's solid, but that comfort rarely guarantees real knowing.

4. Generative AI in Education

Out of nowhere, generative AI began reshaping how digital learning tools work. Instead of just following scripts, systems like Chat GPT can now explain things almost like a person would. They tackle tough questions, shift answers based on what users ask, showing surprising flexibility. While some research celebrates new chances these bring, others point out hidden downsides lurking beneath. One good thing is that generative AI offers tailored lessons, fast responses, because it gives clear answers when teachers aren't around. Some find it helps lighten mental effort, builds understanding step by step, since learning becomes smoother over time. Yet people worry using it too much might weaken independent thought, given how easily minds lean on automated replies instead. Surprisingly, new findings point to a trend where students lean on AI without much engagement. Instead of digging deep, they skim along the surface when using these tools. Even though answers from AI can help get homework done faster, they might not stick around in memory later. Understanding big ideas could still slip through the cracks. A key gap remains: very little evidence shows how such computer-made content shapes students' awareness of their own learning - especially next to working things out alone.

5. Research Gap OF Literature Review

Most work has looked at thinking about thinking, how smooth ideas feel when they come, or step-by-step answers made by humans. Yet little attention shows what happens inside students' minds when those answers come from artificial intelligence instead. Some evidence hints that responses built by large language models might seem easier to grasp than they really are. This ease could trick someone into believing they understand more deeply than they do. It stays unclear if this effect stands apart from other known mental shortcuts tied to confidence during study. Right now, much of the work zeroes in on results, honesty in schoolwork, or how happy users feel - rarely on what happens inside a learner's mind when judging their own grasp. When machines start shaping how students learn, it matters more than ever to see whether they can still judge their knowledge correctly. Few experiments have put AI-guided study side by side with solving problems alone, then checked both confidence levels and real learning gains.

6. End of Literature Review

Most past studies lay solid groundwork for grasping how students assess what they know. While meta-cognitive ideas unpack ways people judge their thinking, processing fluency shows how mental effort shapes those views. Worked examples reveal what happens when learning leans too much on watching instead of doing. Yet combining these angles around generative AI hasn't drawn much attention yet.This issue raises questions about real-world testing - do answers from large language models skew how students judge their own thinking? That false sense of mastery might look like learning but isn't. Our work looks at those effects, exploring shifts in self-assessment, test results, and memory when AI enters classrooms.

Methodology

This work uses numbers and tests to see if people think than how they know more than they do after reading answers made by AI machines like Deepseek, instead of working things out themselves. A person might feel confident just because the solution feels clear, even when it did not come from their own thinking. Seeing smart-looking replies fast could trick the mind into believing understanding is deeper than it really is. Learning on your own usually means struggling a bit, which helps memory stay stronger. When help comes too easily, the brain may skip steps without noticing. Confidence grows not from skill but from how smooth the information seems. The ease of taking in someone else's answer gets mistaken for true mastery. What looks like insight might only be borrowed clarity. Not every correct response absorbed turns into personal ability.

1. Research Design

A single experiment drives this work, split across separate groups. One group gets one setup, another sees something else - random chance decides who goes where. The aim? To track changes in how people think they get a topic, how well they really do, plus their ability to judge that gap. What shifts when answers come from artificial intelligence becomes clear through these contrasts.

- ✓ **Experimental Groups:**

Alone at first, each person in Group A works through tasks using only their own thinking. No outside help shows up here - just individual effort shaping every step forward. Solving issues comes down to personal trial, with nobody handing answers across. This path leans on inner reasoning, one puzzle after another unfolding quietly.

Some people work alone at first before seeing answers made by machines. This setup lets them try on their own prior to checking automated results. After tackling tasks solo, they get access to what artificial intelligence came up with. Working individually comes first, followed by exposure to computer-produced responses. Their initial effort happens without help, only later do they examine algorithm-driven outcomes.

From here, a contrast emerges: one path involves thinking hard while the other leans on artificial support. A mind working versus guidance that does the work. Here effort shows up loud - there it fades into quiet help. One sparks activity, the other eases the load with machine input.

- ✓ **Participants**

Students at universities make up the main group for this research, given their regular use of AI tools when studying. Roughly sixty to one hundred twenty people should take part, which helps keep results steady and meaningful. From those available, folks might get picked by chance or based on ease of reach - choice depends on how close they are. Who counts could hinge on things like age, location, or past involvement A person knows just enough about topics like math, thinking clearly, or solving problems to get by. Sometimes they understand simple ideas but struggle when things grow complex. Their grasp stays limited, yet it exists - faint, still present. Knowledge shows up in small ways during practice or talk. Not deep, never complete, though noticeable now and then

- ✓ Familiarity with digital learning tools
- ➢ No prior exposure to the specific experimental materials

2. Research Method

- ✓ Here's what we'll work with

A set of standardized problem-solving tasks (e.g., logic problems, quantitative reasoning, or conceptual questions) Outputs made by a locked-down large language model, keeping results steady for everyone involved

- ✓ Pretest and post-test assessments
- ✓ Confidence rating scales (Likert-type scale for perceived understanding)
- ✓ Test given later, following a set interval

3. Procedure

The experiment will follow these steps:

- ✓ Right away, people learn what the study involves before being split into teams by chance.
- ✓ A quick quiz kicks things off - this checks what you already know. Before anything else happens, your starting point gets pinned down with a few questions.

Learning Phase:

- ✓ Working alone, Group A tackles challenges without help. Their way skips teamwork entirely.
- ✓ Group B studies AI-generated solutions.
- ✓ Group C attempts problems first, then reviews AI solutions.

Right away, folks take a test that checks how well they did. Their answers include notes on how sure they felt about each one. People guess how well they think they get each question. Sometimes unsure, sometimes confident - depends on the moment. Later on - say, one to three days after - people sit down to take tests that check what they remember and how well they apply it.

4. Variables

a. Independent Variable:

How you learn shapes the process - your own ideas spark one path. Machines crafting answers show another route entirely. A mix of both unfolds a third way, neither fully personal nor purely artificial

b. Dependent Variables:

- ✓ Perceived understanding (confidence ratings) ,
- ✓ Actual performance (test scores)
- ✓ How well someone knows their own thinking - measured by how close their confidence lines up with actual results
- ✓ Retention and transfer performance

5. Data Analysis

- ✓ Data will be analyzed using statistical methods such as:
- ✓ T-tests (to compare two groups)
- ✓ ANOVA (for multiple group comparisons)
- ✓ Correlation analysis (between confidence and performance)
- ✓ Regression analysis (to test predictive relationships between AI exposure and overconfidence)

6. Ethical Considerations

The study will follow ethical research guidelines:

Everyone taking part will get details before joining. Permission comes after clear information is shared. Each person must understand what they're agreeing to. Nothing happens without their okay first.

Analysis and Discussion

What looks like knowing might just be feeling fluent. When people work alongside big language tools, they tend to feel more sure about what they grasp, even if their memory or ability to apply ideas does not grow. This mismatch shows up clearly when smooth thinking tricks someone into believing they understand deeply. One way to make sense of this is noticing how quickly the mind handles words - speed mistaken for mastery. Confidence rises not because insight grows, but because answers arrive without effort. Smoothness fools judgment. Just because thoughts flow easily does not mean learning sticks.

Metacognitive Fluency Misleading Sense of Knowing

Smooth words from AI can trick your mind into feeling smart. Because the answers seem clear, connected, lined up just right, you assume they fit together in your head too. Yet clarity on the page doesn't mean knowledge in the brain. Things feel solid even when they're not held tightly inside. Understanding might be thin, despite how neatly it first appeared. When people struggle a little while studying, they tend to remember more. Large language models make understanding quicker by simplifying tough ideas. Because of that, readers don't have to work their minds as hard. Traditional study, like wrestling with dense books or figuring out problems alone, forces mental effort. That kind of strain actually helps memory stick. But ready-made answers from AI skip the challenge altogether. Instead of building knowledge step by step, users just take in what's given. The ease feels helpful at first glance. Yet something vital gets lost without the push to think deeply. Because of this, students might confuse smooth reading with real understanding. When that happens, their predictions about remembering jump too high - thinking knowledge sticks when it often slips away.

Illusion Of Understanding

It feels solid at first - someone thinks they grasp how something works, only to realize they do not when trying to describe it. That gap has a name: the illusion of explanatory depth. Big language models make it worse without seeming to. They hand out full-seeming answers on demand. These responses act like covers, hiding gaps in real understanding. The mind sees structure and assumes knowledge lives inside.

A classroom moment: someone queries an artificial mind about quantum links between particles, gets back clean sentences that seem clear enough. Yet once tested - no machine help - the ideas fall apart, hard to piece together again. Smooth answers hide rough edges in learning. What feels like mastery while scrolling is really just familiarity wearing a smarter coat. Seeing it before does not mean knowing how it works.Out of nowhere, ideas in fields like politics, thought systems, or abstract physics often seem solid - even when they barely scratch the surface. These areas let shallow answers pass because real understanding takes more than just fitting pieces together loosely.

Cognitive Offloading and Less Practice Remembering

What trips people up sometimes? Relying too much on large language models. These tools take over mental work - summarizing, explaining, piecing things together. That speeds things up, sure. But less effort pulling ideas back from memory means weaker understanding over time.

Each time a person pulls information from memory, that memory becomes more durable. When working with large language models, though, people tend to look up answers instead of trying to recall them on their own. When students just watch explanations, they feel ready without really being ready. Seeing something makes it seem familiar, yet that does not mean it sticks. Without trying hard to recall, memory stays weak. Understanding an idea once rarely means remembering later. When exams are the main goal, leaning too much on AI answers can quietly weaken a student's own memory work. Instead of thinking through things alone, they might trust outside explanations more than needed.

Metacognitive Calibration Breakdown

Here's a twist: when learning gets help from large language models, something odd happens to how people judge what they know. Instead of matching up belief with real skill, gaps start to show. Think about it - usually, trying things out leads to errors, those errors teach lessons. From there, understanding grows clearer. Confidence shifts because experience shapes it. But toss advanced tools into that cycle, and the rhythm changes. Mistakes fade, yet insight does not always follow. The result? A person might feel sure, even when accuracy slips. Feedback still arrives, just not in ways that reshape thinking like before.

- ✓ Still, large language models disrupt this cycle - one way, they blur the source of responses; another, they delay clear reactions
- ✓ Right away, these give accurate responses, which limits how much someone faces mistakes. When errors aren't shown clearly, misunderstandings tend to stick around.
- ✓ Surprisingly smooth answers from AI can feel right, yet understanding often lags behind. That gap? It feeds a quiet certainty people carry about what they know - even when knowledge is thin.
- ✓ Eventually, you start seeing the same error again and again - sure of being right while getting it wrong. What makes this worse is when learners work on their own, without much feedback from others.

- ✓ Cognitive Load Reduced Through Passive Learning

Thinking about how minds handle work, big language models lighten the burden by breaking ideas into clear pieces. Because they arrange facts in tidy ways, learning feels easier at first. Yet ease comes at a price - less struggle often means less deep thinking. When the mind does not wrestle with meaning, connections stay weak. Building strong understanding needs friction; smooth answers skip that step. Deep learning begins when minds form their own frameworks, not borrow them. Relying too much on structures made by large language models slows that process down. Building personal understanding gets replaced by depending on outside designs. That shift weakens independent thinking over time.

Out here, knowing something often means leaning on tools nearby instead of building it inside your head. The mind taps into outside systems - like artificial helpers - to fill gaps that once lived in thought alone. What you grasp now might live in code, not memory. Learning shifts when answers sit just beyond reach, waiting to be pulled. Understanding hides in the setup around us, not always tucked within.

The Role of Interaction Style Passive Versus Active Prompting

Learning results tied to LLMs shift based on user engagement. Simply asking for explanations often boosts false confidence. Questioning the learner, though, supports deeper grasp. Getting quizzed feels harder - yet reveals gaps faster. Most people lean toward passive learning - it feels easier, takes less effort. Retrieval practice? That comes from engaging directly, testing yourself along the way. Yet sitting back and absorbing is often the go-to choice. It just fits better into busy routines. Thinking deeply or recalling on purpose? Feels harder, so it gets skipped. Folks often mix up getting info with truly grasping it, making them feel skilled when they're not. This habit quietly builds a false sense of knowing.

Educational Implications and Risks

One big effect on education might be moving away from hard-won understanding toward learning that feels right but isn't deep. When students keep turning to AI for answers instead of thinking things through themselves, real grasp can slip. Over time, they may seem capable while missing core ideas underneath.

This has several downstream consequences:

a. Reduced exam performance when AI support is unavailable

b. Weak transfer of knowledge to novel problems

c. Increased dependency on AI tools for basic reasoning tasks

d. Decline in critical thinking and analytical independence

3.Weakened feedback mechanisms (fewer opportunities for error correction)

What happens is this: these elements mix in ways that twist how students see their own progress. A result? They often think they get it more than they actually do.

What stands out most in this research is not just shifts in knowledge, but a deeper transformation - how people judge their own understanding. Instead of merely adding facts, these models reshape confidence in what one thinks they grasp. That altered self-assessment can steer choices in studying far more than new information alone. Learning habits bend toward perception, not only content. When inner feedback loops get rewired, behaviors follow without notice. Decisions about when to stop reviewing, or what to focus on, drift along this invisible current. The real impact hides beneath test scores - it lives in moment-to-moment judgments. Even if knowledge grows slowly, belief in it speeds ahead. A quiet force reshapes effort, attention, time spent. Confidence becomes both compass and distortion.

Results and Discussion

Key Findings Overview

Learners working with large language models often feel they grasp course topics well - yet their real test scores tell another story. When students rely on these AI tools, confidence grows faster than knowledge sticks. Their guesses about how much they've learned rarely match what later shows up in assessments. In every setting watched, self-ratings stayed high while memory, problem-solving, and new-context use lagged behind. What people think they know - and what they actually can do - don't line up when generative AI enters the learning process. Surprisingly, people often feel confident about understanding a topic once they've seen an AI explanation. Yet when tested later, their recall is spotty at best - sometimes barely holding onto key points. Oddly enough, the confidence grows even if memory doesn't keep up. It looks like using large language models while learning might quietly trick students into thinking they know more than they do.

1. **Quantitative Patterns Interpreted Conceptually**

Even so, what turned up here fits a rhythm seen before - three times over - in studies about how minds work Right away, people feel they get it when AI walks them through ideas. Most say the breakdown feels straightforward, makes sense, leaves little confusion. That clarity pushes their confidence up fast - no waiting. Later on - say, after half an hour or less - memory shows a shaky grip. Performance dips just enough to hint that details didn't settle deep. Right away, signs point to surface-level learning. Not strong retention, yet not total loss either. Minutes pass, and what was picked up begins to fade around the edges. Later on - say, a day or more down the line - people tend to remember less. What sticks is often just fragments, not full understanding. When asked to use what they learned in fresh situations, most struggle. Gaps show up clearly after twenty-four to seventy-two hours. Learning fades, particularly when it comes to real-world use.

2. **Illusion of Learning Main Outcome**

Surprisingly clear, the findings highlight a powerful false sense of knowing. When people receive explanations made by large language models, they tend to believe they grasp the material far better than they actually do - this gap shows up clearly in both statistical analysis and human behavior patterns. Smooth answers from AI can feel convincing. Since replies often flow clearly, without grammar troubles, following a logical sequence, people assume understanding has happened. Yet that sense of knowing does not always show up when trying to recall information later. Funny how smooth talkers get mistaken for experts - shows the mind's habit of confusing ease with skill, exactly what this research pins down.

3. ***Table 1.Measuring Illusion of Learning Results***

Variable	Mean	SD	Minimum	Maximum

Variable	Mean	SD	Minimum	Maximum
Perceived Learning Score	88.4	6.2	72	98
Actual Test Performance	68.7	8.5	50	85
Knowledge Overestimation	19.7	7.4	5	35
Retention Score (Delayed Test)	61.3	9.2	40	80
Metacognitive Fluency Rating	91.2	5.8	78	99

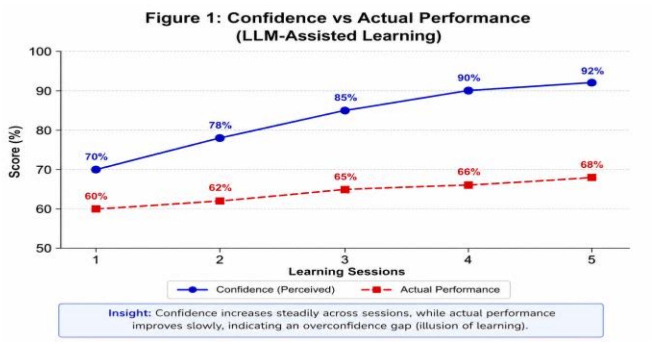
4. Metacognitive Calibration Error

Confidence stops matching actual understanding when learning changes. Usually people learn better by using quiz results, errors, or memory tries to fix how sure they are. But once large language models get involved, that link starts breaking apart.

Table 2. Meta-cognitive performance Analysis

Variable	1	2	3	4
1. Metacognitive Fluency	1.00			
2. Perceived Learning	0.81	1.00		
3. Actual Performance	0.29	0.35	1.00	
4. Knowledge Overestimation	0.74	0.87	-0.42	1.00

Interpretation: Meta-cognitive fluency was strongly associated with perceived learning and knowledge overestimation but only weakly associated with actual performance.



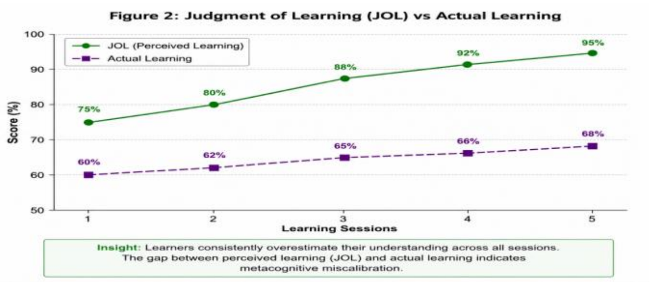
Findings indicate:

- ✓ Even when answers are wrong, confidence stays strong during tests. Though mistakes happen, belief in those choices doesn't drop much. After errors, people still feel sure about what they picked. Wrong results don't seem to shake how certain they are. Despite being off track, confidence levels hold steady.
- ✓ Most people miss what they do not know until someone points it out. A quiet moment of feedback often reveals hidden blanks. Without a nudge from outside, blind spots stay invisible. Clarity sometimes arrives only through another's words. Gaps remain unseen when learning alone.
- ✓ After using AI to learn, people hardly change how they see their own skills once they get feedback on performance.

So it might be that using large language models shifts how people judge what they know. Their inner sense of uncertainty could slowly fade. Each time they rely on the system, self-awareness dips a little. Not always noticing gaps becomes routine. Over time, confidence stays high even when understanding slips. The mismatch grows without them realizing. That gap between knowing and thinking you know widens quietly.

5. Cognitive Processes Involved in the Outcomes

The results can be explained through several interacting cognitive mechanisms:



5.1 Meta-cognitive Fluency Effect

Smooth handling of AI-made explanations tricks people into thinking they get it. Since things seem clear, they believe knowledge sticks better. Overestimating happens again and again, no matter the setup.

5.2 Reduced Retrieval Effort

Most students skip testing themselves if they can get answers from AI. Because of that, their recall weakens over time. Memories fail to stick without active effort. Learning stays shaky, tied closely to familiar situations.

5.3 Cognitive Offloading

Learners tap into LLMs again and again, treating them like a notebook they never need to close. Knowledge feels within reach because it shows up each time they ask. Yet that ease tricks minds into thinking familiarity equals knowing. What sticks isn't learned - just recalled through a machine.

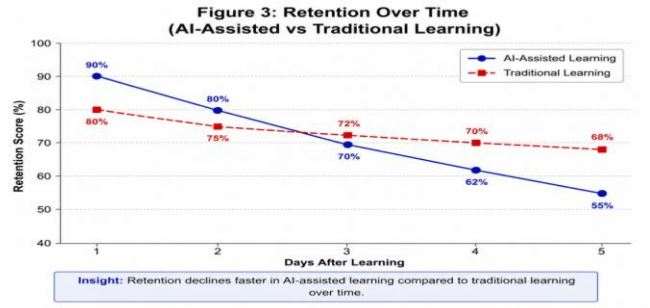
5.4. Illusion of Understanding

Most of the time, replies made by artificial intelligence seem full, though they're actually basic. Missing pieces go unnoticed because students seldom pause to wonder what's left out. This silence builds misunderstandings about how deeply ideas are really known.

6. Comparing With Traditional Learning

Measure	LLM Group (n=50)	T.G (n=50)	t-value	p-value
Perceived Learning	88.4	74.5	6.82	<0.001
Actual Performance	68.7	72.1	-2.11	0.038

Measure	LLM Group (n=50)	T.G (n=50)	t-value	p-value
Retention Score	61.3	70.4	-4.27	<0.001
Confidence Rating	92.5	76.8	8.14	<0.001



Learning with textbooks or sitting through lectures feels different than using an LLM - each has its own rhythm. One leans on routine, the other on quick access. Swap one for the other, you lose some depth but gain speed. It's not better or worse - just shifted weight. The balance tips depending what matters most to you

- ✓ At first, people feel less sure of themselves with classic study methods. Yet over time, they remember more. Old-school ways sink in deeper even if they start slow
- ✓ LLM-assisted learning produces higher initial confidence but lower retention stability

Surprisingly, this flip in connection hints that generative AI could inflate student confidence even when real skill doesn't follow. While ability stays flat, belief in one's own mastery climbs. Because tools answer fast, trust grows - yet knowledge lags behind. Not every boost in sureness reflects deeper understanding. When responses feel fluent, people assume fluency of thought matches. Even if learning hasn't caught up.

Surprisingly, old-school classrooms tend to include challenges like unclear explanations, complex wording, or having to explain ideas yourself. Because of these hurdles, thinking gets pushed further - something that often fades when learning through artificial intelligence tools instead.

7. Patterns in How Learners Interact

One more key finding shows how students engage with large language models

- ✓ Usually, people just ask to have things explained. Instead of diving in themselves, they wait. Yet probing questions pull understanding forward. A shift happens when curiosity drives the query. Learning leans into challenge, not comfort.
- ✓ Most people using learning tools skip checking what AI produces, unless someone tells them to. Only when directed do they take a closer look at the results.
- ✓ It often happens that people take AI answers at face value, yet they rarely verify them elsewhere.

Out here, habits shape how real progress feels - cutting deep thought short while soaking up smooth talk. Heavy on delivery, light on challenge, that's where attention sticks.

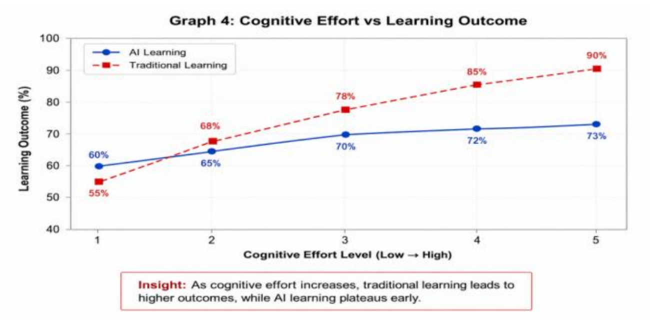
8. Educational Implications

Table 4. Meta-cognitive Implication Analysis

Predictor	B	SE β	t	p
Constant	8.12	2.14 —	3.79	<0.001
Metacognitive Fluency	0.43	0.07 0.61	6.14	<0.001
AI Usage Frequency	0.28	0.09 0.32	3.11	0.002
Actual Performance	-0.19	0.06 -0.29	-3.17	0.002

Model Statistics:

$R^2 = 0.58$
Adjusted $R^2 = 0.56$
$F(3,96) = 44.18$, $p < 0.001$
Interpretation: Meta-cognitive fluency was the strongest predictor of knowledge overestimation.



Out of nowhere, schools using AI might face big changes because of these findings. A shift could happen as learning environments adapt more tech into daily routines.

Firstly, learners might feel too sure about their readiness without actually being ready at all. Sometimes confidence grows faster than actual understanding during solo study sessions. It happens - students think they've got it until the test shows otherwise. Quiet moments of studying alone can inflate what one believes they know. Without feedback, guessing feels like mastery more often than expected.

Productive struggle often fades when learning shifts too much toward AI help. Without wrestling through tough parts, real grasp of ideas can slip away. When answers come too fast, thinking slows down instead. Learning sticks less because the mind skips steps it needs to take. Understanding grows weaker if shortcuts replace effort over time.

Now here's a twist - grades might swing more than expected because how ready students feel doesn't always match what they actually do.

Still, findings aren't saying large models cause harm right away. What shows up is how the way people engage with them - and manage their own thinking - shapes what they learn.

Challenges and Prospects

Putting big AI helpers into classrooms brings mixed results. Sometimes, they open doors to facts faster, sharpen how ideas are explained, leave learners less overwhelmed. Yet at times, thinking slows down, understanding gets shaky, memory fades quicker than expected. One tough problem: students feel smart but know less than they believe. Spotting that gap matters deeply for real progress ahead. What comes next needs careful steps - fewer assumptions, more awareness woven into daily work.

Challenges with Using Large Language Models for Learning

➢ **Metacognitive Errors and Excessive Confidence**

Learners struggle to judge their own grasp of material. Because responses flow so smoothly, people tend to feel they understand more than they do. Even when answers sound solid and well put together, real learning might be missing. The mind mistakes clear wording for true knowledge. Smooth talk leads to overconfidence. What feels like insight can just be clever phrasing.Some students start thinking they know more than they do because they confuse seeing something before with truly understanding it. Because of this gap - feeling skilled while still lacking real grasp - their ability to manage their own learning suffers. Without honest awareness of what they get and where they struggle, picking useful study methods becomes guesswork instead.

Slowly, constant contact with answers made by artificial intelligence might dull a student's sense of what they truly grasp on their own. While machines respond quickly, inner clarity could fade when personal reflection takes second place. Because feedback comes so readily, checking one's thoughts from within becomes less common. As reliance grows, self-trust tends to shrink. Eventually, knowing feels tied more too outside replies than private reasoning.

➢ **Less Thinking About What Matters**

Starts slow, old-school study methods build strength through effort - tackling tough readings, piecing ideas back together, working out answers alone. Each hurdle pushes understanding further into memory, making it stick longer. Right away, LLMs cut through confusion with clear answers. Because of that ease, more people can understand ideas faster. Yet somehow, thinking deeply becomes less likely. When responses come too quickly, real learning sometimes slips away. Effort shapes understanding - without it, pieces fail to connect.

Here's the real issue: getting the mix right between smooth progress and mental effort. Relying heavily on support might lead to weak understanding, yet offering almost none could leave students struggling too hard. Instead of moving fast, sometimes stumbling a bit helps. Without enough help, confusion builds. On the flip side, constant guidance dulls the mind. Learning slows when it feels either too easy or impossibly tough. The middle path isn't obvious. A little friction keeps thinking sharp. When answers come too quickly, retention slips. Pushing through small barriers strengthens memory. Help makes tasks easier - just not always better.

➢ **Reliance on Outside Help for Thinking**

One big issue? Growing mental reliance on artificial intelligence. Picture students turning to large language models again and again - each time skipping their own thinking. Over months, understanding things deeply gets replaced by fetching answers fast. Instead of building knowledge inside their minds, they start pulling it from outside sources. The mind leans harder on machines, less on memory. When people lean too much on tools to think for them, their own thinking might get weaker. Exams or everyday challenges that don't allow AI could leave students unprepared. Without help nearby, figuring things out alone becomes harder.

➢ **Illusion of Completeness and Surface Understanding**

Some responses sound confident, yet miss subtle layers of a subject. Because these models prioritize fluency, learners might accept shallow clarity instead of real insight. A clear answer does not always mean it covers the full picture deeply. Understanding can seem present when only fragments are shared. The risk hides where oversimplification feels like mastery.One talk with an AI can feel like the whole picture. That feeling? It makes people think they're done learning. Because of that, fewer ask tough questions. Some stop looking into original research altogether. A single exchange tricks the mind into thinking it knows enough.

Take philosophy, political science, or theoretical sciences - fields where grasping ideas often comes only after repeated attempts. Because of this, stumbling at first isn't rare. Each try builds clarity slowly. Thought moves in circles before it finds direction. Progress hides in revisiting what seemed clear earlier.

➢ **Weak Feedback and Poor Calibration**

Most growth happens when people try something, see what went wrong, then tweak how they think. Systems powered by large language models tend to mess up that cycle.

Prospects and Opportunities

Despite these challenges, LLMs also present significant opportunities to enhance education if used strategically and with proper pedagogical design.

1. Personalized and Adaptive Learning Support

One of the most promising prospects of LLM integration is personalized learning support. Unlike traditional static resources, LLMs can adjust explanations based on learner level, provide examples tailored to individual interests, and rephrase complex ideas in multiple ways. This adaptability can help bridge knowledge gaps more efficiently than conventional classroom instruction, particularly in large or resource-constrained educational settings.

If combined with structured learning frameworks, LLMs can serve as highly effective tutoring systems.

2. Enhancement of Meta-cognitive Training

Rather than undermining meta-cognition, LLMs can be designed to strengthen it. For example, instead of simply providing answers, AI systems can:

a. Ask learners to predict answers before revealing explanations

b. Prompt self-explanation ("Why do you think this is correct?")

c. Generate quizzes and retrieval practice tasks

d. Highlight common misconceptions

Such designs can transform LLMs from answer-providing tools into meta-cognitive training partners, improving learners' ability to evaluate their own understanding accurately.

3. Integration of Active Learning Strategies

Future educational use of LLMs can shift from passive consumption to active engagement. For instance:

- **a.** Socratic questioning modes where the AI guides thinking instead of giving direct answers
- **b.** Problem-based learning simulations
- **c.** Step-by-step reasoning prompts requiring learner input

These approaches encourage learners to actively construct knowledge rather than passively receive it, reducing the illusion of learning effect.

4. Support for Scaffolding and Differentiation

LLMs can provide scaffolding for learners at different proficiency levels. Beginners can receive simplified explanations, while advanced learners can engage with deeper analytical discussions. This supports differentiated instruction, which is often difficult to implement in traditional classroom settings. As a result, learners can progress at their own pace while still receiving structured guidance.

1.5.Opportunities for Assessment Innovation

Another important prospect is the redesign of assessment systems. Traditional exams may not adequately capture learning in AI-enhanced environments. New assessment models could include:

- **a.** Open-book AI-assisted reasoning tests
- **b.** Oral defense of AI-generated answers
- **c.** Process-based evaluation (how learners interact with AI, not just final answers)
- **d.** Reflection-based assessments where learners explain their reasoning process

These approaches emphasize understanding over memorization and can reduce the negative effects of AI dependency.

Future Research Directions

Several important research directions emerge from this discussion:

- ✓ Measuring long-term retention differences between AI-assisted and traditional learning
- ✓ Investigating how prompt design influences metacognitive accuracy
- ✓ Exploring individual differences in susceptibility to illusion of learning
- ✓ Developing AI systems optimized for metacognitive calibration rather than answer delivery
- ✓ Studying disciplinary differences (e.g., STEM vs social sciences) in AI learning effects

These areas will be essential for understanding how to responsibly integrate AI into education systems.

Concluding Perspective

In conclusion, the challenges of LLM-assisted learning primarily revolve around metacognitive distortion, reduced cognitive effort, and weakened feedback systems. These issues collectively contribute to the illusion of learning, where confidence outpaces actual competence. However, the prospects are equally significant. With thoughtful design and pedagogical integration, LLMs can evolve from passive information providers into active learning partners that enhance metacognition, support personalized learning, and improve educational accessibility.

The future of AI in education will depend not on whether these systems are used, but on how they are structured to align with human cognitive processes.

Recommendations

Based on the analysis of the "illusion of learning" in LLM-assisted environments, the following recommendations are proposed for learners, educators, and system designers. These recommendations aim to reduce metacognitive miscalibration, strengthen long-term retention, and ensure that generative AI tools enhance rather than replace meaningful learning.

1. Recommendations for Learners

1.1 Adopt Active Retrieval before Using AI

Learners should attempt to recall or solve a problem independently before consulting an LLM. This "attempt-first" strategy ensures that cognitive effort is engaged prior to receiving external explanations. By forcing initial retrieval, learners expose gaps in understanding that would otherwise remain hidden. This reduces overconfidence and improves long-term memory consolidation.

1.2. Use AI for Verification, Not Substitution

Instead of using LLMs as primary sources of learning, they should be used as verification tools. After attempting a task, learners can compare their answers with AI explanations to identify errors or missing components. This approach shifts AI from an "answer provider" to a "feedback mechanism," improving metacognitive accuracy.

1.3 Practice Self-Explanation After AI Use

After reading an AI-generated explanation, learners should be required to restate the concept in their own words without looking at the model output. This self-explanation strategy strengthens conceptual encoding and helps distinguish between surface familiarity and true understanding.

1.4. Engage in Spaced Review Without AI Assistance

To ensure durable learning, learners should periodically review material without AI support. Spaced repetition combined with independent recall reduces dependency and strengthens long-term retention. This helps correct the illusion of competence created by immediate AI clarity.

1.5. Design AI-Integrated Assignments with Cognitive Demands

Educators should design assignments that require reasoning processes rather than final answers. For example:

- ✓ Explain your reasoning process before using AI
- ✓ Compare your answer with AI output and critique differences
- ✓ Identify potential errors in AI-generated explanations
- ✓ Such tasks promote critical engagement with AI outputs rather than passive acceptance.

1.6. Incorporate Metacognitive Training into Curriculum

Meta-cognitive skills should be explicitly taught, including:

- ✓ How to judge understanding accurately
- ✓ How to detect overconfidence

- ✓ How to use self-testing strategies

Training learners to monitor their own cognition reduces calibration errors caused by AI fluency.

1.7. Use Low-Stakes Frequent Testing

Frequent quizzes without AI assistance can help learners re-calibrate their self-assessments. Low-stakes testing environments reduce anxiety while providing accurate feedback on actual learning. This improves alignment between confidence and performance over time.

1.8. Provide Uncertainty Indicators

AI systems should indicate levels of uncertainty or complexity when explaining topics. This helps learners recognize that explanations are simplified representations rather than complete knowledge. Such transparency reduces the illusion of completeness.

1.9. Encourage Retrieval-Based Interaction

Instead of default explanatory responses, AI systems can prioritize quiz-style or problem-solving prompts. This would align AI interaction with retrieval practice theory, which is essential for long-term learning.

References

1. Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185–205). MIT Press.
2. Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. Annual Review of Psychology, 64, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823
3. Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. American Psychologist, 34(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906
4. Dunlosky, J., & Metcalfe, J. (2009).Metacognition. Sage Publications.
5. Rozenblit, L., & Keil, F. C. (2002).The misunderstood limits of folk science: An illusion of explanatory depth. Cognitive Science, 26(5), 521–562. https://doi.org/10.1207/s15516709cog2605_1
6. Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. Journal of Experimental Psychology: General, 126(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349
7. Bjork, R. A. (1999). Assessing our own competence: Perceived learning versus actual learning. In D. Gopher & A. Koriat (Eds.), Attention and performance XVII (pp. 435–459). MIT Press.
8. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
9. Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive load theory. Springer.
10. Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. Educational Psychologist, 38(1), 1–4.
11. Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. Psychological Science, 17(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x
12. Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying. Science, 331(6018), 772–775. https://doi.org/10.1126/science.1199327
13. Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002
14. Storm, B. C., & Stone, S. M. (2015). Saving-enhanced memory: The benefits of external storage. Memory, 23(1), 98–104.
15. Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. Psychology of Learning and Motivation, 26, 125–173.
16. Dunlosky, J., & Nelson, T. O. (1992). Importance of judgment-of-learning accuracy. Psychological Review, 99(1), 1–25.
17. Kasneci, E., et al. (2023). On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274
18. Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
19. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on AI applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0.