FALSE COGNITIVE POWER TRANSFER

From Individual Atrophy to Collective Demoralization in the Age of AI

Author: Adrian (Adi) Stan
Independent researcher, Pitesti, Romania
ORCID: 0009-0003-1457-5155

ABSTRACT

Generative artificial intelligence produces a cognitive paradox: it amplifies the capabilities of experts, but generates the dangerous illusion of competence in inexperienced users.

This article introduces the concept of **False Cognitive Power Transfer (FCPT)** - the phenomenon whereby individuals mistakenly attribute the effectiveness of AI outputs to their own cognitive competence, leading to taking on more tasks or responsibilities than they can realistically, systematically, manage.

The FCPT triggers two collective failure mechanisms:

1. **Collective Demoralization Cascade (*CDC*)** - spectacular failures lead to collective preemptive abandonment.
2. **The Replication Illusion (*RI*)** - observing successes generates overconfidence in replicability, ignoring that AI functions as a cognitive exoskeleton that amplifies existing skills and knowledge, but does not create them.

In the long term, both mechanisms lead to **cognitive atrophy** - the progressive loss of the original deep-thinking capacity, even in experts. The article proposes the ***Mechanism of Cognitive Stimulation* (*MCS*)** as a countermeasure intervention.

CHAPTER 1. THE PROBLEM

Generative AI has democratized access to sophisticated cognitive capabilities. A student generates code in unfamiliar languages. A manager produces strategic analyses without a consulting background. A junior researcher formulates complex experimental hypotheses without decades of experience.

This apparent democratization masks a critical problem: **the confusion between tool efficiency and user competence.** When AI produces quality output, the user experiences **FCPT** - the illusion that the cognitive power in the output comes from their own mental capabilities.

The existing literature documents individual cognitive atrophy - GPS reduces spatial orientation, computers diminish mental arithmetic skills. But the focus remains on individual skill loss, not the social consequences of the illusion of competence generated by tools.

For years, studies have shown us that GPS makes us lose our sense of direction, and computers weaken our mental calculation skills. But these studies talk about personal atrophy. What they don't explain is how the illusion of competence turns into collective demoralization and social blockage.

This mechanism - from individual illusion to collective blockage - is exactly what is missing from the public discussion. We are not talking about how AI makes us "stupid". We are talking about how we trick ourselves into thinking we are smarter than we are, and then misinterpret the social signals around us, creating a spiral of distrust. **Why is this happening?** Because in the absence of clear and immediate feedback telling us „Stop, this is your limit", our brains take a dangerous shortcut: **they confuse the result with the process.** See a good result = assume you have the process internalized. But in fact, AI keeps the process in the "black box", and you just press the button.

There is a lack of analysis of the intermediary mechanisms through which AI generates negative effects even in the absence of malignant intent:

- How access to powerful tools generates systematic confusion between the tool's capability and the user's capability.
- How this confusion can lead to unrealistic commitments in complex projects.
- How the resulting spectacular failures are socially interpreted as evidence of the impossibility of the task, not of erroneous estimation.
- How this interpretation generates collective abandonment of important works, perfectly achievable with correctly calibrated expectations.
- Why observing the successes of others doesn't protect us, but makes us believe that it's easy for us too, leading to our own failures.

This analysis is not a critique of AI. It is a diagnosis of how, when we are not careful, **technological amplification can degrade our ability to properly assess ourselves and learn from experience - both individually and collectively** .

CHAPTER 2. THE FUNDAMENTAL ANALOGY: The Carpenter and the Jackhammer

An experienced carpenter works with a hand hammer. He can roof a small house in 10 days. He gets a Jackhammer - now he roofs the same house in 3 days.

REAL power transfer: the hammer amplifies existing muscle strength. After a period of constant use, the hand muscle atrophies, and the carpenter even forgets how to hammer nails by hand. Up to this point, the phenomenon is documented in the literature under the name "cognitive atrophy through neglect." This happens because both our brain and body are efficient - we deactivate what we don't use. When the tool does everything, we gradually disconnect from the basic process.

The critical problem arises when the same carpenter declares, *“I can now cover the Cathedral in ten days.”* If the task were simply to hammer in nails, the estimate might be plausible. But the Cathedral requires much more: complex architecture, coordination between crafts, historical understanding, specialized materials, layered aesthetic decisions. In simple terms: the house is a **problem of volume** (more nails), while the Cathedral is a **problem of type** - it requires a completely different kind of thinking and experience.

After a month of effort, the carpenter declares defeat, the roof is damaged, and the project fails spectacularly.

The most serious consequence is NOT the individual failure of the carpenter, but the collective reaction: the other carpenters - with or without a Jackhammer - conclude: *"If the Cathedral Carpenter was not able to build the roof with a Jackhammer, we certainly cannot!"*

The result: The Cathedral remains roofless. Not from a lack of real capabilities, but from an erroneous recalibration of collective trust based on the observation of a failure generated by unrealistic expectations. This is the essence **of the Cathedral Problem:** powerful instruments generate not only individual atrophy, but also a collapse of collective trust in the possibility of carrying out complex works.

It's what I call **the "Reverse Pavlov Syndrome".** If Pavlov trained the dog to salivate at the light, the success of the experiment "trained" Pavlov to turn on the light. In the age of AI, we think we are using a tool, but its speed trains us not to think. We enjoy the "food" (response) received instantly, without seeing that we have become **addicted to the button that turns on the light** .

On the other hand, an apprentice, seeing the carpenter with the jackhammer who finished the roof of the house in 3 days, thinks: *"With that jackhammer, I can roof a house in 3 days too!"* The result? Apprentice + jackhammer = (still) insufficient. The apprentice fails. The problem is NOT the lack of the jackhammer (which exists), but **the erroneous attribution** - the belief that the "power" comes from it, and not from the carpenter's experience, amplified by the tool. The key idea is that the tool is neutral - it amplifies whatever is given to it. Giving it to an expert amplifies expertise. Giving it to a novice amplifies only the illusion of competence.

In modern terms: Writing a routine email with AI is like roofing a small house - it's efficient and secure. But writing the security architecture of a bank (the Cathedral) using AI, without being a security expert yourself, is a recipe for disaster. The code will look perfect, run without syntax errors, but it will have open doors that only a human architect would have guessed.

CHAPTER 3: False Cognitive Power Transfer (FCPT)

Artificial intelligence tools are the contemporary equivalent of the jackhammer - but for cognitive work, not physical.

False Cognitive Power Transfer (FCPT) is the psychological mechanism by which the user of a powerful tool systematically confuses the tool's effectiveness in specific tasks with his or her own ability to manage complexity, make strategic decisions, and orchestrate multidimensional processes.

FCPT components:

1. **Demonstrable efficiency** - The tool generates real productivity gains in well-defined tasks. Roofing a house can indeed be done in three days with the Jackhammer.
2. **Cognitive atrophy** - Constant use of the tool diminishes fundamental capabilities. The carpenter forgets how to hammer nails by hand. The analyst forgets how to structure problems strategically without AI. The phenomenon is extensively documented in the literature.
3. **Categorical Confusion (new element)** - The user unjustifiably extrapolates from speed in simple tasks to capacity in complex projects. Erroneous logic: "If I can cover a house in three days, I can cover the Cathedral in ten days." But the Cathedral is not a "multiplied

house", but a fundamentally different object - it requires architecture, coordination, history, aesthetics.

4. **Structural hubris** - Categorical confusion generates unrealistic commitments. The user overestimates, not out of ignorance (Dunning-Kruger), but from extrapolation valid in a limited domain, erroneously applied to extended domains.

FCPT vs. Dunning-Kruger Effect

Dunning-Kruger effect explains how **incompetence generates overestimation through the inability to recognize one's own ignorance.** But this mechanism is INTERNAL - it results from a lack of metacognitive knowledge, not from access to powerful external tools.

FCPT describes an EXTERNAL ATTRIBUTION error, induced by interaction with a technological system that amplifies performance.

- ***Dunning-Kruger*** *example:* Piano student thinks he can play Chopin after 3 months -> tries -> quickly realizes he can't.
- ***FCPT*** *Example:* Student uses AI for sophisticated musical interpretation of Chopin -> excellent output -> believes he understands at an expert level -> has difficulty reproducing without AI support.

The critical difference: In Dunning-Kruger, the error disappears with learning. In FCPT, the error can INCREASE with assisted experience because the tool masks the lack of competence.

Why is it important to distinguish FCPT from simple laziness or Dunning-Kruger ? Imagine a GPS: the first time you use it, you know you wouldn't have known the way. The second time, you look at the map less. The tenth time, you don't even remember the names of the streets you've traveled. This isn't laziness - it's a "rewire" of the brain. Your brain learns that "navigation is externalized" and simply deactivates its spatial orientation modules. FCPT does exactly the same thing, but with the "complex problem-solving" module. You don't become lazy, you become incompetent without realizing it.

Why "categorical confusion"? Because the human brain has a system error: if you succeed in the "simple tasks" category, mental automatisms tell it that you will also succeed in the "complex tasks" category. But the categories are fundamentally different - just as "running 100m" and "running a marathon" are not the same category of activity. The AI makes you run 100m in 3 seconds. Your brain deduces: "Then the marathon will be fast too." The error occurs because you do not perceive that the marathon requires strategy, endurance, energy management - things that you have never practiced because the AI only delivered you sprints.

Structural hubris vs. personal arrogance: This is not about arrogant people. It is about the system that creates arrogance. The manager does not wake up in the morning thinking “I am the smartest”. He sees that the AI delivers 10 reports a day, so logically he accepts to lead a strategic project. The company praises him for his speed. When the project fails, the system (the company, the market, the colleagues) does not say “you overestimated”, it says “the project was too risky”. Thus, the structure (pressure for speed, praise for output, lack of critical feedback) generates the hubris, not the person. That is why it is “structural”, not personal.

CHAPTER 4: Empirical Examples (2021-2025)

Zillow Offers (2021)

- **Impact:** Losses of over $500 million / Closure: November 2021 / Staff reductions: approximately 2,000 employees
- **Root cause:** Real estate valuation algorithms that overestimated property values.
- **Fundamental error:** Confusing local predictive performance with the ability to manage real estate market dynamics.
- **CEO statement (Rich Barton):** "The unpredictability of the market has far exceeded what we had anticipated."
- **FCPT:** The belief that a performing model can substitute human judgment in complex systemic contexts.

Olive AI (2023)

- **Impact:** Losses of approximately $832 million / Peak valuation: $4 billion (2021) / Layoffs: 450 employees (July 2022) + 215 (February 2023)
- **Cause:** The promise of near-total automation of medical processes, without an adequate understanding of the operational constraints in healthcare.
- **Fundamental error:** Confusing point automation with systemic transformation.
- **FCPT:** Overestimating AI's ability to replace human expertise in a critical and hyper-regulated field.

Builder.ai (2025)

- **Impact:** Losses of over $450 million / Maximum valuation: $1.5 billion / Status: Insolvency
- **Root cause:** Operating model based on human labor disguised as full automation. Details: Approximately 700 human developers supported what was presented as an “AI platform.”
- **Structural issues: Overstated revenues,** unsustainable operating costs. Debts: $88 million to AWS, $30 million to Microsoft.
- **FCPT:** Customers and investors confused the promise of automation with the actual existence of a scalable solution.

Why aren't these examples just "companies that screwed up"? Each of these cases illustrates a different deception of the FCPT, and cumulatively they show how versatile and dangerous this mechanism is.

Zillow: The fallacy of absolute prediction. When you look at these numbers, understand the context: Zillow was not just any company. It had the best real estate data in the US, thousands of engineers, and billions of dollars. If they thought they could predict house prices better than human agents, it wasn't out of arrogance, but because their algorithm actually performed great in tests. FCPT infiltrated the system: the entire company interpreted success in 100,000 transactions as proof that they could scale to millions, ignoring that the real complexity was not in the prediction, but in managing market uncertainty - something that a human agent with 20 years of experience does instinctively, but AI does not. When the market became volatile (COVID, inflation), the AI continued to provide accurate but irrelevant predictions - and the company lost half a billion believing them.

Olive AI: The Deception of Total Automation. Healthcare is the most complex human system: rules, ethics, patient variability, bureaucracy, human life. Olive promised hospitals: “We will automate 90% of processes.” What investors didn’t see: they didn’t automate anything, they just interfaced robots over fragile human procedures. When one case went out of pattern, everything collapsed. Here, FCPT manifested itself on two levels: internally: management thought that because AI could process a form, it could manage an entire hospital, and externally: customers thought that because Olive had a $4 billion valuation, it “works.” Both sides confused speed of execution with understanding the system. Result: $832 million evaporated and hundreds of people laid off.

Builder.ai: The AI product deception versus the disguised human service. This is the clearest case of the IR (Replication Illusion) of all. Builder.ai promised: “Tell us what app you want; the AI will build it automatically.” Reality: Behind the scenes, 700 Indian programmers worked, writing the code manually. The AI was just a “translator” between the client and the team. But the text on the website, the investor pitch, the demos - everything suggested complete automation. Clients and investors saw the output (fast-delivered apps) and attributed the success to the AI, not the hybrid human-AI process. When the real costs became unbearable (700 salaries versus the promise of zero scaling), the model collapsed. Lesson: If you can’t accurately say who did what (human vs. machine), you’re in full FCPT. And the consequence is bankruptcy.

Why does all this matter to you? Because **you’re not immune.** If you use ChatGPT to write emails and feel increasingly confident in your communication skills, you’re on the same path as Zillow. If you see a friend who makes an AI app in 3 months and think “I can do it too”, you’re the Jackhammer apprentice. If you’re a manager and you delegate a strategic project to a team that “has AI”, without checking to see if they understand the domain - you’re the investor in Builder.ai. These cases aren’t about technology, they’re about attribution. And misattribution costs maybe half a billion dollars. Every time.

These failures show that we are hitting a hard **Psychological Ceiling. When the difference** between how smart you are and how smart the AI seems **Once you cross a certain threshold,** your mind stops cooperating - it either surrenders (L1) or goes on the defensive. That's why we see that huge gap in the economy: some fly with AI (L3), others collapse under it (L1/L2).

CHAPTER 5: The Cascade of Collective Demoralization (CDC)

False Cognitive Power Transfer does not remain isolated. It generates a cascade of systemic errors, through which local failures turn into structural blockages.

Typical sequence:

1. **Initial failure** - an actor overestimates the capability of the AI-assisted system.
2. **Social observation** - failure becomes visible to other actors.
3. **Misattribution** - the cause of failure is attributed to fundamental domain limitations, not to evaluation error.

4. **Defensive generalization** - actors avoid similar projects, considering them impossible to achieve.
5. **Systemic blockage** - viable initiatives are abandoned preemptively.

In many situations, the final result is therefore not predominantly determined by the lack of technical capacity, but by the loss of collective confidence in the possibility of building complex systems.

The Collective Demoralization Cascade (CDC) is a social phenomenon whereby the visible failure of an actor, occurring in the context of the use of advanced technologies, generates the preemptive abandonment of similar initiatives, probably often achievable under the right conditions. The central element of this process is not the failure itself, but the erroneous attribution of the cause of the failure: what is, in reality, an implementation or evaluation error is reinterpreted as a structural limitation of the respective domain.

CDC mechanism:

1. **Phase 1: Initial Failure** - The individual ("carpenter") with FCPT undertakes complex project ("Cathedral"), visibly fails, generating damage ("damaged roof").
2. **Phase 2: Social Observational Learning** - Observers do not experience failure themselves, but observe it in others. Social psychology (Bandura, 1977) shows that observational learning is as powerful as direct experience in shaping behavior.
3. **Phase 3: Misattribution** - Observers attribute failure to "impossibility of task even with powerful tool", not to "misestimation of complexity". Why? Because: the tool was obviously powerful (house covered in three days), the individual seemed competent (successful in simple tasks), the failure was categorical (unfinished Cathedral). Natural attribution: *"The task is beyond human capabilities, even technologically augmented."*
4. **Phase 4: Preemptive Collective Abandonment** - The other "carpenters" - including the competent ones who could realistically complete the Cathedral roof - preemptively give up: *"If X with a Jackhammer couldn't do it, I certainly can't."*
5. **Phase 5: The cathedral without a roof** - The important work remains unfinished not due to a lack of real capacity (there were competent carpenters), but due to a collapse of collective trust generated by a socially amplified individual failure.

Historical case studies:

1. **Case 1: Cold Fusion (1989–present)** - The 1989 announcement of the achievement of nuclear fusion at low temperatures generated massive interest. Later, the difficulty of replicating the results led to the conclusion that the phenomenon was physically impossible. Consequences: almost complete reduction of government funding for research in the field, stigmatization of the term "cold fusion" fusion" in academia, the migration of research towards alternative terminologies (LENR). Although subsequent studies indicated reproducible anomalies under certain conditions, the field remained marginalized for over three decades. It was only after 2020 that significant funding programs reappeared ($10M ARPA-E in 2023).
2. **Case 2: Gene Therapy (1999-2013) - In 1999,** Jesse 's death Gelsinger during a clinical trial triggered a severe institutional backlash. Funding was drastically reduced, and the field entered a decade-long decline. Although subsequent investigations indicated problems specific to the vectors used and safety procedures, the prevailing perception

was that gene therapy as a whole was too risky. It was only after the development of safer vectors (AAV) and revision of protocols that the field began to recover, culminating in clinical approvals after 2017.

3. **Case 3: The AI "Winter"** - In the 1970s and 1980s, the over-promises of symbolic AI led to major disappointments. The Lighthill Report (1973) and the failures of expert systems triggered massive funding cuts. The result was a decade-long period of stagnation, known as the "AI Winter", in which the field was marginalized, despite the existence of viable technical directions. Only with the emergence of new paradigms (machine learning, big data, hardware acceleration) was the field rehabilitated.

Why aren't these just "technical failures" but failures of interpretation? Look at the pattern: in each case, the technology worked - Zillow 's algorithms predicted prices, Olive processed the form, Builder took applications. The failure wasn't technical, it was epistemic - that is, it was a cognitive error, not a code error. The people involved didn't know what they didn't know. They didn't know that predicting local prices doesn't mean understanding the market. They didn't know that processing a form doesn't mean curing a patient. They didn't know that fast delivery doesn't mean automation. And when the system crashed, this interpretation error was invisible - everyone just saw "the technology failed."

Why is observational learning so dangerous here? Normally, when you see someone fall on ice, you learn to be careful. But in the case of the CDC, you see someone fall with professional climbing gear. The conclusion is not "he didn't know how to use it", but "you can't climb the mountain with that kind of gear." Why? Because your brain uses a mental shortcut that works 99% of the time: if a high-performance tool fails at a visible task, the task must be impossible. This is a survival fallacy - in nature, if the strongest lion can't catch the buffalo, it gives up. But in modern society, this shortcut makes us abandon perfectly achievable cathedrals just because someone used the jackhammer incorrectly.

History repeats itself - and it's no coincidence. Cold Fusion, Gene Therapy, AI Winter - they all have the same CDC pattern. Why? Because our society develops technologies exponentially, but our social learning mechanisms are linear. We learn from failures through observation, but we observe too little and generalize too quickly. In a year, an AI company can generate more spectacular failures than all corporations generated in the previous decade. And each failure is not learned in isolation - it propagates instantly through social media, news, podcasts. Result: the cascade of demoralization has accelerated 1000x in the AI era. Where in the past a failure in gene therapy blocked the field for 15 years, today a Zillow failure can block investments in real estate AI for the next financial cycle.

Why do the most competent give up first? The CDC paradox is that the most capable actors are the most vulnerable to demoralization. Why? Because they have a clearer picture of the real complexity of the Cathedral. When they see someone fail, they think: "If he, with all the resources of the AI, did not succeed, then the project is really impossible." The less competent, on the other hand, are protected by their illusion (RI) - they believe that they can succeed. Thus, **the CDC removes** exactly **the most valuable actors** from the game and **replaces them with overconfident amateurs.** The system is reversed: **specialists give up, dilettantes try.** The result is a reverse selection in innovation - the market is filled with unviable projects, and those who could have built real Cathedrals sit on the sidelines.

CHAPTER 6: THE REPLICATION ILLUSION (RI)

The Replication Illusion (RI) is the phenomenon whereby the success achieved by an actor through the use of artificial intelligence systems is interpreted by observers as being easy to reproduce, often without taking into account their actual level of competence.

Why is it important to distinguish RI from FCPT? RI occurs after success, not after failure. FCPT makes you believe that you are competent. RI makes others believe that your success is easy to copy. RI is contagious, it spreads through social networks, pitches, LinkedIn. FCPT is individual. This is not a theory, but an observation of a pattern that repeats itself in hundreds of cases.

The difference between **understanding** and **copying:** you can perfectly understand how a washing machine works, but that doesn't mean you can build it. Why is RI more dangerous than FCPT? Because FCPT only **destroys the individual, RI It destroys** the entire **ecosystem.** RI creates a market of people who think they can do something they can't, investors who fund unviable projects, and companies that lose money replicating models that don't work.

How does RI feel in real life? Imagine seeing a post on LinkedIn: *"I built an AI app for legal contracts. $100,000 monthly revenue in 6 months."* You think: *"I can do that too!".* RI just got activated. That guy has maybe 10 years of legal experience - he knows what clauses matter. The AI just sped up his writing. You only see the result and the tool, but the invisible expertise that makes success valuable doesn't show up in the post.

RI is the virus behind the "tutorials" that don't work. Why don't you become an expert by watching a tutorial ? Because the instructor (**L3 expert**) implicitly knows what questions to ask, when to ignore the AI, how to adapt. You copy the steps and get a similar result. RI activates: *"I did the same, so I know!".* But you only copied the surface. The real expertise - knowing what happens when the AI makes a mistake - remains invisible and, crucially, uncopied.

Why does RI destroy the ecosystem, not just the individual? FCPT makes you think you are smart. RI makes others think they are smart too, creating a cascade of incompetent imitation: one startup succeeds → RI generates 100 superficial imitations → 90% fails → investors don't see the failure of RI, but "AI doesn't work" → trust drops → legitimate projects suffer. RI doesn't just create amateurs – it eliminates professionals.

The washing machine analogy - why isn't it enough? So, you can understand how a washing machine works without building it. AI adds a danger: you watch a "How it works" video and think you can open a factory. AI makes you jump from conceptual understanding (I know the theory) to executional ability (I can build) while ignoring 99 practical sub-skills. Even worse: **you don't know you don't have them,** because the AI has delivered the finished result to you.

RI and adverse selection in innovation. RI creates a **perverse paradox:** the most competent actors (L3) see the wave of imitations and give up building Cathedrals because the market is choked with weak huts. In return, the least competent (L1 with RI) enter en masse, believing they can build. Result: adverse selection - the market is filled with unviable products, and legitimate projects no longer receive funding. RI does not just create imitations - it eliminates the original.

The antidote to RI: Transparency of real effort. The solution is not "more tutorials", but the deliberate exposure of invisible effort: the real learning time, the number of failures, the hidden costs. If every successful post included "7,200 hours of prompt engineering", RI would

be disabled. But social platforms suppress this information because it decreases engagement. Thus, RI is algorithmically amplified: what is viral is what is superficial, and what is superficial creates the illusion of replication.

How to protect yourself from IR as an observer. When you see the next post “I did *xyz* with AI in 3 days”, ask yourself: What am I missing? I’m missing maybe 10 years of experience, or maybe 1,000 previous failures, or the ability to evaluate what’s good. IR makes you see only the tip of the iceberg. The solution is to force awareness of the whole iceberg - to ask yourself questions about what you’re not being shown. Only then can you distinguish between replicable success (simple procedures) and expert success (which is not copied, it’s earned).

CHAPTER 7: COGNITIVE ATROPHY - The Delayed Cascade Effect

Cognitive atrophy can be defined as a progressive diminution of the capacity for deep, original and sustained thinking, as a delayed effect of dependence on AI and learning from failures generated by FCPT, CDC and RI.

The critical difference from FCPT: FCPT is an illusion of competence in the present (you believe you are capable now), and Atrophy means real loss of competence in the future (you become progressively incapable).

Causal mechanism: How cognitive atrophy sets in:

- **Year 1:** FCPT -> "I am competent with AI"
- **Year 2:** Failures (CDC + RI) -> "AI is not good enough"
- **Year 3:** Increased dependency -> "Why try without AI?"
- **Year 5:** Giving up on your own effort -> "You'd do better anyway"
- **Year 10:** Atrophy -> significantly reduced capacity without AI (in certain types of tasks)

Wrong learning after repeated failures generated by FCPT/RI:

- **Correct learning:** "I overestimated; I should have assessed my abilities better."
- **Mislearning (common):** "The problem is that the AI isn't strong enough."
- Mislearning -> increased dependency -> accelerated atrophy.

What does "cognitive atrophy" actually mean in real life? You don't wake up one morning and forget how to think. No. It's much more subtle. Let's say you work in marketing and you use AI to generate posts. At first, you use AI for inspiration, but you rewrite them. After 6 months, you help the AI edit them. After 2 years, you can't write a post from scratch without feeling "stuck." That's atrophy: the mental muscle that generates original ideas has atrophied from disuse. You're not less intelligent - you're just less able to access that intelligence without the AI shortcut. Just like the carpenter who forgets how to hammer nails by hand.

Why the “delayed cascade effect”? Because it’s not immediately visible. FCPT tells you “You are competent” today. Atrophy tells you “You are no longer competent” only in 3 years. Why? Because the loss of ability is gradual and invisible as long as the AI is present. You only discover the atrophy when you need to think for yourself - and you can’t. When the AI is not available, when the context is new, when the problem is atypical. Then you realize: you confused

access to the solution with **the ability** to solve. I personally went through this. Fortunately - there is an antidote.

"Year 1, Year 2, Year 3, Year 5, Year 10" timeline - what does it look like in reality? Let's take the example of a programming student:

- **Year 1 (FCPT):** Use ChatGPT for homework. Get a high grade. Think: "I'm good at Python !" (Illusion).
- **Year 2 (Failures):** Receives a complex project - ChatGPT delivers code that doesn't solve the problem. Thinks: "AI isn't advanced enough for what I need." (Misguided learning).
- **Year 3 (Dependence):** Accepts that “all good code comes from AI.” Stops trying to write from scratch. Justifies: “Why waste time when AI can do it more efficiently?” (Effort abandonment).
- **Year 5 (Installed Atrophy):** When asked to write a simple function without AI, he gets stuck. He doesn't know where to start. He says, "Wait, I don't have access to ChatGPT." The mental muscle of the original generation has atrophied.
- **Year 10 (Deep Atrophy):** Becomes senior. Leads teams. But when he has to architect a new, complex system, completely dependent on AI - he can no longer evaluate the architecture without AI suggestions. He risks making bad strategic decisions because he no longer has the cognitive muscle to feel what is right.

Why is “mislearning” so common? Because it’s comfortable and logical at first glance. You fail with AI. The natural conclusion: “I need better AI.” Not: “I need to assess my skills better.” Why? Because acknowledging your own limitations hurts. It’s a cognitive conflict: “I’m competent, but I failed. How can that be?” The AI provides a “ scapegoat”: “It wasn’t me who was weak, the AI was weak.” This reasoning protects the ego in the short term, but accelerates atrophy in the long term - because you’re not addressing the real problem.

How is atrophy different from FCPT in your professional life? FCPT is the moment when you present an AI-generated report and feel like an expert. Atrophy is the moment (2 years later) when your boss asks you: “And what do you think?” - and you don’t have an opinion, because you haven’t thought critically about the subject in years. FCPT is the illusion. Atrophy is the bill you pay when the illusion disappears.

Why does atrophy accelerate? Because it's a positive (feedback) cascade. The more you depend on AI, the less you practice independent thinking. The less you practice, the more difficult it becomes to think independently. The more difficult it becomes, the more you prefer to use AI. The loop is self-amplifying. The only stopping point is deliberate awareness - but that only comes if you're exposed to situations without AI that show you that you've lost something. Without that, the atrophy is completely invisible until Year 10.

The most serious problem - Atrophy at the L3 level ("Architects")

Why is it more critical at the expert level:

L3 Expert:

- It has advanced original thinking capabilities, built through thousands of hours of deliberate practice.

- It may gradually reduce its cognitive involvement through dependence on AI-generated patterns.
- Perceives AI-assisted output as "good enough", which masks the gradual loss of evaluative finesse.
- The main cost is the accumulated erosion of expertise developed over the years.

Specific mechanism at L3:

- **Before AI integration (≈2020):** The expert generates multiple candidate solutions, evaluates each option in depth (subtle criteria, accumulated experience), selects the optimal solution based on calibrated intuition.
- **In the first years of AI use (2024-2025):** AI generates a large number of candidate solutions; the expert evaluates and selects the optimal one. The result is often superior to that obtained without AI. The process represents a form of healthy augmentation.
- **After 5+ years (2029-2030):** AI generates most solutions, the expert tends to choose the first good enough option. The in-depth evaluation process is practiced less frequently. A gradual decrease in the ability to make fine discriminations is observed.
- **After 10+ years:** The ability to generate original solutions without AI support is significantly reduced. Deep assessment skills become more difficult to access without external support. A form of functional dependence on assistive tools emerges.

Result: A gradual transition from autonomous expertise to mediated competence, in which performance remains high, but cognitive autonomy progressively diminishes.

Why is L3 atrophy catastrophic? Imagine a ship architect with 30 years of experience, who uses AI for structural calculations. At first, the AI generates 10 variants, and the architect critically evaluates them and chooses the best one - perfect augmentation. After 5 years, the AI generates 3 variants, and the architect quickly chooses the first one that seems sufficient because he "knows" that the AI is optimized. After 10 years, when he has to design a new type of ship without precedent, he can no longer generate even an initial variant by himself - he has forgotten how to start from scratch. This is the hidden cost: the AI has preserved his performance, but has stolen his autonomy to think when there is no pattern. When a real crisis occurs (a ship breaks down, a project fails), he no longer has the ability to diagnose because he has not done this deep evaluation for years. Case performance remained high (operations in 2 hours instead of 4), but decision-making autonomy atrophied - and when the AI made a mistake, the architect no longer had the calibrated intuition to correct it. This is **the danger of L3 atrophy:** you are not weaker; you are more vulnerable to the rare events in which the AI makes a mistake. Because you have not practiced thinking for yourself in 1,000 previous cases.

Why are experts perhaps the most vulnerable? Because they have the most to lose and the least incentive to test their limits. A junior knows they don't know, so they test. An L3 expert stops testing because they "know". When the AI delivers an answer, the expert quickly evaluates it and accepts it - they trust their own judgment. But that judgment is calibrated over millions of hours without the AI. Over time, the calibration deflates because the AI is doing most of the evaluation. The expert doesn't realize that his intuition is degrading because he no longer has raw, authentic feedback. He becomes like a veteran pilot flying only on autopilot - in an emergency, his reactions are those of a novice.

Why does “good enough” kill expertise? Because expertise is not about *enough,* it is about *optimal.* An L3 architect does not choose the “good enough” strength structure, but the one that resists earthquake, corrosion, wear, with minimal cost. This finesse of evaluation is achieved by practicing hard selection - not quick choice. When AI generates 3 variants and you choose the first good enough one, you give up the exercise of fine discrimination. This exercise is exactly what makes the expert an expert. Without it, the discrimination muscle atrophies. After 10 years, you can no longer tell the difference between good and excellent - and the cost is hidden in the structure of the ship that will break in 20 years.

What does L3 atrophy look like in medicine? Imagine a cardiologist surgeon with 15 years of experience using AI for diagnosis. The AI suggests: "Urgent surgery." The surgeon evaluates, confirms. The AI suggests: "Low risk." The surgeon accepts. After 5 years, the AI suggests a non-standard surgical technique. The surgeon, accustomed to accepting the suggestion, applies it. But he no longer has the ability to assess whether that technique is appropriate for the patient's specific anatomy - because he hasn't done this in-depth assessment in years. Result: post-operative complications. Case performance remained high (operations in 2 hours instead of 4), but decision-making autonomy atrophied - and when the AI was wrong, the surgeon no longer had the calibrated intuition to correct it. This is the danger of L3 atrophy: you are not weaker; you are more vulnerable to the rare events in which the AI is wrong. Because you have not practiced thinking for yourself in 1,000 previous cases.

Why is L3 atrophy invisible until the moment of crisis? Because until Year 10, performance is perfect. You evaluate AI reports quickly, you fix AI code superficially, you approve AI decisions because they are “well-founded.” No one around sees the degradation - because the output is of quality. The atrophy is not in the output, it is in the process of generation. When AI is no longer there, the process is the missing airbag. When an unprecedented situation arises, you no longer have the mental modeling to solve it - because AI has delivered you ready-made models. The crisis reveals the atrophy, but until then there are no external signs. That is why it is catastrophic: it is detected too late, and the repair (re-calibration of intuition) requires years of deliberate effort - exactly what you have been avoiding for the last decade.

CHAPTER 8: MECHANISM OF COGNITIVE STIMULATION (MCS)

Mechanism of Cognitive Stimulation prevents cognitive atrophy by deliberately introducing "useful friction" - calculated delays and challenges that force the user to process information actively, not passively.

Central principle: When AI instantly answers complex questions, the user does not have to go through the cognitive journey required for deep learning. MCS introduces calibrated interventions that restore cognitive effort.

The design challenge: How to introduce friction without frustrating the user? The answer: dynamically adapting to the individual cognitive trajectory.

Think of MCS not as a limitation, but as a Gym for the Mind. When you go to the gym, you don't want the barbell to lift itself (that would be zero friction). You want the weight to provide resistance, because **resistance builds muscle** .

Likewise, a healthy AI shouldn't give you the "bite-in-the-mouth" answer to strategic problems. It should act like a sparring coach. partner): to challenge your ideas, ask for arguments, and force you to think, not just edit.

How MCS works:

- **MCS 1.0** is the basic protocol: If the complexity of the question is high and the AI answers too quickly, the system introduces artificial latency OR asks a clarifying question. Purpose: Forces the user to process the information.
  MCS levels: Level 0: Direct response (without prompting); **Level 1:** Clarification ("Is the wedding in the evening or during the day?"); **Level 2:** Synthesis ("What is the priority: speed or security?").
- **MCS 2.0** introduces **Extended Levels** and bidirectional feedback: it detects whether the user is progressing (ascending), regressing (descending), exploring laterally, or stagnating - and dynamically adjusts the challenges. This way, users who are actively progressing are given more complex challenges, and those who are struggling are given reduced cognitive pressure.

What does "useful friction" mean in real life? Think of a fitness trainer: if they give you a weight that's too light, your muscles won't grow. If they give you a weight that's too heavy, you'll get hurt. MCS is the brain's personal trainer. When the AI responds instantly, it's like someone lifting the weight for you - your brain isn't doing any work, it's just assisting. MCS puts the weight down and says, "You lift it, but I'll give you 3 extra seconds to get ready." That's the calibrated challenge.

Why isn't artificial latency just "slowing down for the sake of slowing down"? Latency is not punishment, but space for reflection. When the AI doesn't respond instantly, your brain doesn't sit idle - it automatically starts looking for the answer on its own. You may not find the whole solution, but you activate your neural networks that make connections, test hypotheses, look for patterns. When the AI then offers the answer, you don't receive it as a tablet from heaven, but as confirmation or correction of what you already thought. This is active learning - the only one that prevents atrophy.

How does MCS 2.0 detect whether you are "progressing" or "regressing"? MCS 2.0 does not read thoughts, but question patterns. If you go from "How do I do X?" to "But what happens if I modify X like this?" - the system sees semantic construction (i.e. you use concepts from the previous answer). This means progress. If you go to "Excuse me, can you repeat?" or "I don't understand, explain it more simply" - MCS detects regression or blockage. Then, it automatically increases the friction threshold - it gives you a direct answer so as not to frustrate you, but notes that you need practice.

Why is dynamic adaptation crucial? A rigid protocol (MCS 1.0) would be like a trainer who gives you the same weight whether you’re a beginner or a champion. You’d frustrate the beginner and under-stimulate the expert. MCS 2.0 is the personal trainer who sees your face - if you wince in pain, reduce the weight; if you make the exercise too easy, add more. In technical

terms: μS (the measure of cognitive demands) is not static. It goes up when you're tired, it goes down when you're fit, and MCS 2.0 tracks it.

How does a "Level 2: Synthesis" work in practice? Let's say you ask: "How do I increase sales?" AI 1.0 would instantly respond with 10 tactics. MCS Level 2 first asks: "What's the priority: speed or security?" - forcing you to clarify the strategic context, not just look for tactical solutions. When you answer "speed", the system knows to offer aggressive growth tactics, not consolidation ones. But the cognitive effort to decide the priority was yours, not the AI's. That means you did the hard work - and your brain stayed trained.

Why is MCS 2.0 essential for education and corporations? In schools, teachers cannot sit next to each student to see when they are "progressing" or "stagnant". MCS 2.0 can. If a student asks increasingly sophisticated questions about a topic, the system can give them bonus challenges (Level 2-challenge). If another student repeats the same question, the system can intervene with a human tutor - not continue to give unhelpful answers. In corporations, MCS 2.0 can diagnose teams - if in a department all employees are regressing (constantly asking for clarification, not synthesizing), management knows that they are addicted to AI and need training. MCS becomes not just a coach, but a cognitive audit tool.

The future is not about who has the fastest AI (that battle is already lost by humans), but about who knows how to adjust their " phase". The proposed solution, which I call **MSC 3.0,** will be a partnership in which you and the AI will openly negotiate: " Hey, for this problem I need to be more creative, will you let me increase the processing temperature?" **This is symbiosis: when both, the human and the AI, will have their "hand" on the intelligence tap.**

Without protection = technology becomes dangerous not because of what it is, but because of how it is used.

ARE YOU IN DANGER? The quick 3-question test

Before you say "this doesn't happen to me," answer honestly:

1. **The Panic Test:** If OpenAI's servers go down tomorrow for a week, can you finish your current project at the same quality (albeit slower), or are you completely stuck?
2. **Verification Test:** When the AI gives you a complex answer, do you spend more time checking the logic or formatting the text? (If you're just formatting, you're in the risk zone).
3. **Learning Test:** After solving a problem with AI, could you explain "why" the solution worked to a junior colleague, without looking at the history?

CHAPTER 9. CONCLUSIONS

Generative artificial intelligence creates a fundamental paradox: while amplifying the capabilities of experts, it simultaneously generates dangerous illusions of competence in inexperienced users and risks progressive cognitive atrophy even in experts.

Contributions of this article:

1. Defining **FCPT** as a psychological mechanism distinct from Dunning-Kruger, whereby attribution confusion between tool effectiveness and one's own competence generates systematic over-commitment to tasks of inappropriate complexity.

2. Identification of two patterns of collective failure post-FCPT: **CDC** (Negative learning from observed failures -> collective preemptive abandonment) and **RI** (Distorted positive learning from successes -> ignoring expertise -> failure to replicate).
3. Explaining **cognitive atrophy** as a delayed cascade effect, arguing that degradation can affect not only beginners but also L3 experts, through progressive dependence on AI patterns and the erosion of original deep-thinking capacities.
4. **MCS 2.0** (or **3.0**) proposal as an adaptive pyramidal protocol of counteraction through useful friction dynamically calibrated to the individual cognitive trajectory.
5. Presentation of empirical evidence from multiple fields (Cold Fusion, Gene Therapy, AI Winter, Zillow, Olive AI, Builder.ai) suggesting that CDC and RI patterns are observable and have recurring historical analogies.

Central message: AI is neither salvation nor apocalypse. It is a **cognitive exoskeleton** - it amplifies what already exists, it does not create from nothing.

Our responsibility is threefold:

A. **Awareness:** We recognize **FCPT**, **CDC**, and **RI** when they occur.
B. **Responsible design:** We implement **MCS** 2.0-type mechanisms.
C. **Sustained effort:** We deliberately maintain fundamental cognitive capabilities.

The forked future:

- **Trajectory 1 (dystopian):** Society cognitively dependent on AI, widespread loss of fundamental capabilities, severe vulnerability when AI systems fail or change, gap between those who retain capabilities (elite) and those who atrophy (the majority).
- **Trajectory 2 (healthy augmentation):** AI as a conscious amplification partner, deliberate maintenance of capabilities through MCS, education recalibrated for the AI era, responsible design of systems with anti-atrophy safeguards.

We are in the early years of the AI era. The design decisions, educational policies, and social norms established over the next **3-5 years** will determine which trajectory we follow.

The alternative to conscious action is not the status quo - it is the drift towards Trajectory 1. Cognitive atrophy does not require malignant intent. It only requires **inaction** .

Technology replaces skills - that's **evolution.** The critical question is not "Do I use AI?", but "Do I control what I lose?". How many people still know how to mow with a scythe? Almost none - and that's OK, because it's no longer necessary. But structuring thinking, critical evaluation, generating original solutions - these cannot be outsourced without cost. **The difference** between **evolution** and **addiction** is **conscious control:** you know what you delegated, why, and **you can come back when necessary.** Even the farmer who can cut himself in the tractor plow - the next day, after LEARNING the lesson, continued to work with the tractor. He didn't go back to the scythe. And he didn't dismantle the tractor either. But he uses protective gloves when changing the plow. And, so that he doesn't **forget** to mow, he uses the scythe where the tractor can't reach the field, or maybe just in the garden in front of the house.

Rogo, ergo emergo.

(I question, therefore I become.)