# Critical Thinking and AI in Education: Unpacking the Domain-Specificity Question

## The Problem with "High Critical Thinking Students Benefit More from AI"

Claims that students "high in critical thinking" profit more from learning with AI likely conflate several distinct constructs. The domain-specificity question is central to understanding what's actually happening—and has significant implications for how we design AI-integrated learning environments.

---

## The Domain-General vs. Domain-Specific Debate

### The Traditional View (Domain-General)

The traditional assumption has been that critical thinking is a transferable skill—that once learned, it can be applied across any domain. This view underpins many "21st-century skills" initiatives and generic critical thinking courses.

### The Cognitive Science Challenge (Domain-Specific)

Daniel Willingham's work, drawing on decades of cognitive science research, argues that "critical thinking (as well as scientific thinking and other domain-based thinking) is not a skill. There is not a set of critical thinking skills that can be acquired and deployed regardless of context."

The processes of thinking are fundamentally intertwined with content knowledge. This creates what Willingham calls a "surprising failure to deploy useful knowledge" across contexts.

**Key evidence:**

- Experimental evidence shows that experts don't think as well outside their area of expertise, even in closely related domains

- Knowledge of medicine transfers poorly among subspecialties—neurologists do not diagnose cardiac cases well

- Technical writers cannot effectively write newspaper articles

- Even professional philosophers are swayed by irrelevant features of problems like question order or wording

**The Practical Implication**

Abstract principles like "look for hidden assumptions" won't help much in evaluating an argument about a topic you know little about. More fundamentally, if you lack background knowledge about the topic, you likely won't even comprehend the claims being made—writers and speakers routinely omit information they assume their audience already knows.

---

## What Studies Actually Measure When They Measure "Critical Thinking"

Studies measuring "critical thinking" in AI contexts are likely capturing one or more of the following constructs:

| Construct | Description | Relationship to "Critical Thinking" |
|---|---|---|
| **Domain expertise** | Deep knowledge within a specific field | Enables what *appears* to be critical evaluation of AI outputs |
| **Prior academic achievement** | Historical performance measures | Correlated with both domain knowledge and general cognitive ability |
| **Metacognitive strategies** | Planning, monitoring, evaluation skills | Somewhat generalizable, but require domain knowledge to execute |
| **General cognitive ability** | g-factor, working memory, processing speed | Correlates with performance across domains but isn't transferable "critical thinking" |

The conflation problem: When a study finds that "high critical thinking students" benefit more from AI, it may actually be measuring any combination of these constructs.

---

## The Expertise Mechanism vs. the "Critical Thinking" Mechanism

### What's Likely Happening

When "high critical thinking" students benefit more from AI, I hypothesize this is actually an **expertise effect**:

- A biomedical science expert can spot when ChatGPT gets biochemistry wrong

- A novice cannot make this evaluation regardless of their "critical thinking skills"

- This looks like "better critical thinking" but is really "more domain knowledge enabling quality evaluation"

### Evidence from AI Education Research

Research on AI-assisted critical evaluation in biomedical science found that high-performing students "might be better equipped to navigate or compensate for the limitations of the AI tool in understanding complex academic topics." This reflects expertise, not transferable critical thinking.

Similarly, research on intelligent tutoring systems shows that effectiveness varies based on prior knowledge—but this is because prior knowledge enables meaningful engagement with the material, not because "critical thinking" transfers.

---

## The Metacognition Question

### A More Defensible Position

Certain **metacognitive strategies** are somewhat generalizable:

- Planning how to approach a task

- Monitoring comprehension during learning

- Evaluating one's own understanding

- Recognizing when to seek help

These strategies can be learned and applied across contexts. However, **executing** these strategies still requires domain knowledge.

### The Execution Gap

A student might learn the metacognitive prompt "check if this sounds plausible" and attempt to apply it across contexts. But:

- "Check if this biochemistry is plausible" requires knowing biochemistry

- "Check if this historical claim is plausible" requires knowing history

- "Verify AI outputs against authoritative sources" requires knowing what authoritative sources exist in that domain

The metacognitive strategy without domain knowledge is an empty instruction.

### Applied to AI

A student might learn "verify AI outputs against authoritative sources"—a legitimate metacognitive strategy. But verification requires:

1. Knowing what authoritative sources exist in the domain

2. Understanding the domain well enough to evaluate consistency

3. Recognizing what kinds of errors AI tends to make in that domain

4. Having sufficient background knowledge to notice omissions

Research shows that "people without research backgrounds tend to overly rely on AI to answer questions." This isn't a failure of critical thinking instruction—it's a consequence of lacking the domain knowledge that would enable critical evaluation.

---

## Implications for AI in Education

### If Critical Thinking Were Domain-General

**Prescription:** Train students in general critical thinking skills, then they'll be able to use AI well across contexts. Focus on "AI literacy" as a transferable competency.

### If Critical Thinking Is Domain-Specific (The Evidence-Based View)

**Prescription:** Students need to develop genuine expertise within domains before AI becomes a beneficial tool rather than a crutch. General "AI literacy" training won't substitute for learning actual content.

This aligns with:

- **Cognitive Load Theory's expertise reversal effect:** Scaffolding appropriate for novices may harm advancing learners

- **Desirable difficulties research:** Productive struggle builds the knowledge structures that enable later evaluation

- **The calculator/GPS research:** Tools that help experts may harm novices still building foundational competence

---

## Practical Recommendations

### For Course Design

1. **Sequence matters:** Students likely need foundational domain knowledge before AI assistance becomes beneficial rather than harmful

2. **Don't assume transfer:** "AI literacy" skills learned in one domain may not transfer to others

3. **Build evaluation capacity through content:** The ability to evaluate AI outputs comes from domain expertise, not generic critical thinking training

### For Research

1. **Disaggregate constructs:** Studies should measure domain knowledge, general cognitive ability, and metacognitive strategies separately

2. **Test transfer claims:** Does "critical AI use" learned in one domain transfer to novel domains?

3. **Examine the mechanism:** When "high critical thinking" students benefit more, is this actually expertise?

**For Teaching AI Literacy**

1. **Embed in domains:** Teach AI evaluation within specific subject areas, not as a standalone skill

2. **Be skeptical of transfer claims:** Don't assume students who learn to evaluate AI in one context can do so in another

3. **Prioritize content knowledge:** The best "AI literacy" training may be ensuring students have the domain knowledge to evaluate AI outputs

---

## Key Takeaways

1. **Critical thinking is not a transferable skill** in the way commonly assumed—it's deeply embedded in domain knowledge

2. **"High critical thinking" in AI studies likely measures domain expertise**, not a generalizable evaluation capacity

3. **Metacognitive strategies are somewhat generalizable** but require domain knowledge to execute

4. **Educational implications differ dramatically** depending on which view is correct—and the evidence favors domain-specificity

5. **The best preparation for critical AI use may be deep learning within domains**, not generic "AI literacy" training

---

## Key References

- Willingham, D. T. (2008). Critical Thinking: Why Is It So Hard to Teach? *Arts Education Policy Review*, 109(4), 21-32.

- Willingham, D. T. (2019). How to Teach Critical Thinking. *NSW Department of Education Occasional Paper Series*.

- Willingham, D. T. (2020). How Can Educators Teach Critical Thinking? *American Educator*, Fall 2020.

- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612-637.

- Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, 22, 123-138.

- Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications. *Instructional Science*, 38(3), 209-215.

---

*Document prepared for research on AI and cognitive offloading in higher education*