

Literature Review: The Intersection of AI and Self-Regulated Learning

Conceptual and Theoretical Foundations of SRL

Self-Regulated Learning (SRL) is defined as a crucial higher-order skill wherein learners actively and constructively set goals, monitor progress, and regulate their cognition, motivation, and behavior to achieve learning objectives. The development of SRL competencies is particularly critical in online learning environments, where learners often experience a scarcity of traditional tutor guidance.

Several influential theoretical models underpin research in this domain. Zimmerman's cyclical model is frequently adopted as a framework for designing and evaluating AI tools. This model organizes SRL into three phases: Forethought (task analysis, goal setting, motivation), Performance (self-control, self-observation/monitoring), and Self-Reflection (self-judgment, self-reaction). Similarly, Pintrich's framework delineates four phases, emphasizing four areas of regulation: cognition, motivation/affect, behavior, and context. While SRL models generally form a coherent and integrative framework, interventions based on these models have differential effects based on the learner's educational level; for instance, models grounded in socio-cognitive theory (like Zimmerman's) are generally more impactful in primary education, while those focusing on metacognitive aspects (like Winne and Hadwin's) are often more beneficial in secondary settings.

AI Applications and Support for SRL Processes

The literature documents a surge in research on AI and SRL, especially following the emergence of generative AI (GenAI) around 2022. This research primarily targets the higher education level and predominantly utilizes three types of AI applications: chatbots, Intelligent Tutoring Systems (ITS), and AI-powered evaluation systems. The majority of studies focus on the engineering and basic science disciplines.

AI tools are strategically integrated to support the core phases of SRL:

- **Forethought (Planning/Goal Setting):** Historically, goal setting was one of the most prevalent processes supported by SRL tools. AI chatbots continue this function by acting as personalized planners and goal-setting facilitators, enabling tasks such as defining objectives and receiving strategy recommendations.
- **Performance (Execution/Monitoring):** AI applications frequently target the performance phase. This includes systems providing real-time task scaffolding and feedback to assist with monitoring and actively tracking progress. For

example, studies in programming education analyzed interactions where students mainly utilized AI as a reactive debugging tool for code correction and error interpretation (Monitoring phase).

- **Reflection (Evaluation/Adjustment):** Chatbots act as metacognitive scaffolds by prompting learners to critically evaluate their strategies and outcomes. Specific implementations, like the customized GPT model in a neuroanatomy course, successfully leveraged conversational AI to foster metacognitive awareness and readiness for self-directed learning (SDL) through regular reflective practice. The design of SRLAgent, for instance, explicitly roots its gamified environment in all three phases of Zimmerman's model to enhance overall SRL skills.

The Dual Impact of AI: Support, Risks, and Design Implications

The integration of AI presents a paradox: it offers tremendous support but simultaneously introduces significant risks to the learning process.

Positive Outcomes and Acceptance

A significant majority (69%) of empirical research reports a positive impact of AI support on SRL outcomes. The introduction of generative AI often leads to measurable improvements in performance; for example, the ChatGPT group in one study significantly outperformed both human expert and control groups in improving essay scores. Furthermore, acceptance levels are generally high among pre-service teachers, driven by factors like perceived AI usefulness, perceived AI trust, and perceived AI enjoyment. Successful generative AI designs demonstrate effectiveness not merely in providing answers but in intentionally supporting personal development and requisite skill development for lifelong learning.

Risks and Cognitive Deterioration

The most critical challenge identified across multiple sources is the phenomenon of metacognitive laziness (or cognitive offloading). This occurs when learners delegate necessary cognitive tasks to AI tools, such as ChatGPT, circumventing the mental effort required for deep learning, problem-solving, and critical self-regulation. This offloading can lead to superficial understanding, short-term performance gains that mask long-term skill stagnation, and diminished independent learning skills. Supporting this risk, neuroscientific research indicates that extensive reliance on LLM-generated content is associated with diminished neural activity in brain regions linked to semantic integration and executive self-monitoring.

The technical characteristics of LLMs introduce specific pedagogical risks, including:

- Model-level issues: Hallucinations, algorithmic bias, and privacy concerns.
- Behavioral risks: Over-reliance, diminished critical thinking, and reduced neural activity.

Intentional Design and Scaffolding Strategies

To mitigate these risks, designers must ensure AI functions as a cognitive amplifier that complements human intelligence rather than replacing essential processes. This requires intentional design strategies, including:

Metacognitive Scaffolding Designs must actively promote critical self-evaluation and learner agency (Habermas's emancipatory interest). Tools should employ indirect scaffolding, such as hints, step-by-step plans, and Socratic questioning, rather than directly providing final solutions.

Role Inversion (Cognitive Mirror) The proposed "Cognitive Mirror paradigm" advocates for inverting the traditional "AI as Oracle" role. Here, AI acts as a teachable novice designed to reflect the quality of the learner's explanation, forcing the learner to actively structure knowledge and confront misconceptions. This is achieved by repurposing AI safety guardrails (Diversion Guardrail Mechanism) to deliberately limit the AI's knowledge scope, creating a "pedagogically useful deficit".

Heutagogy The AI-augmented heutagogical design framework posits that AI should be integrated to foster self-determined learning and learner agency across progressive paradigms (AI-Informed, AI-Supported, AI-Empowered). This approach emphasizes double-loop reflection and active documentation (e.g., AI-use reflection forms) as safeguards against over-reliance.

Convergence and Divergence in the Literature

Convergence (Shared Consensus and Reinforcement)

The Dominant Threat of Cognitive Offloading The most consistent finding across contexts (general education, programming, critical thinking) is the pervasive risk of excessive reliance on AI leading to cognitive offloading and "metacognitive laziness". This risk mandates that AI integration prioritize engagement over efficiency.

Focus of SRL Support Both older reviews on tool design and newer studies on AI integration identify goal setting, monitoring, and self-evaluation as the key SRL processes that technological tools aim to support.

AI for Short-Term Performance AI/LLMs demonstrate a significant capability to boost short-term, task-specific performance (e.g., essay quality or debugging speed).

Need for Intentional Scaffolding There is a consensus that merely deploying powerful AI is insufficient; systems require explicit pedagogical safeguards, such as indirect prompts, hints, or restricted functionality, to encourage critical reasoning and overcome the tendency toward passive acceptance.

Relevance of Core SRL Models Zimmerman's three-phase model (Forethought, Performance, Reflection) is the most frequently adopted theoretical framework guiding the design and analysis of AI-supported SRL interventions.

Divergence (Contrasting Findings and Interpretations)

Motivation vs. Performance Outcomes One randomized study found that while the ChatGPT group achieved significantly better essay scores than all other groups (including human experts), there were no significant differences in intrinsic motivation, knowledge gain, or knowledge transfer. This suggests that performance improvements might be transactional (optimizing for the rubric) rather than rooted in genuine motivational engagement or deep learning. Conversely, other studies note that AI tools can positively influence motivation and perceived enjoyment, which contributes to higher acceptance.

Effectiveness of Metacognitive Dialogue While some research demonstrates that custom generative AI successfully fosters metacognitive awareness and readiness for SDL, experimental studies using LLM-powered chatbots for reflection reported low student engagement, brief interactions, and no evidence that increased engagement leads to improved academic performance. This highlights a tension between the theoretical potential and the practical efficacy of using chatbots solely for metacognitive prompting.

Optimal Design Paradigms The literature presents contrasting high-level design philosophies: the "AI as Oracle" approach (rich content generation, solving problems efficiently, which underlies much ITS and checklist tool design) versus the "Cognitive Mirror" paradigm (inverting the role to force active knowledge construction by the learner). The former risks metacognitive laziness, while the latter intentionally introduces desirable difficulties to foster deep learning.

Integration Focus within the SRL Cycle While general SRL tool reviews highlight goal setting, monitoring, and self-evaluation as most supported, observational studies of AI usage (e.g., programming assistants) show that student interaction overwhelmingly centers on the Monitoring phase (reactive error-fixing). This reveals a gap between planned pedagogical support and actual student-driven behavior with the tools.

Overall Synthesis and Future Design Implications

The literature confirms that AI and LLMs represent a critical turning point for SRL research and practice, offering unparalleled opportunities for personalization, instant feedback, and scaffolding across the learning cycle. The central synthesis, however, lies in navigating the tension between efficiency and intellectual effort. The immediate performance gains provided by AI are frequently decoupled from deeper indicators of learning, such as motivation, knowledge transfer, and metacognitive skill development.

Successful integration demands a move beyond merely creating AI support tools towards establishing a principled pedagogical design architecture that actively counters cognitive offloading. Future systems must embrace complexity by incorporating not only the technical interest (efficiency, personalization) but also the practical interest (social co-regulation via teacher-peer-chatbot triads) and the emancipatory interest (critical reflection, ethical inquiry, and fostering learner agency). Conceptual frameworks like the AI-augmented heutagogy and the Cognitive Mirror paradigm offer actionable roadmaps for designing AI systems that intentionally transform the learner's role from passive recipient to active, responsible knowledge constructor, ultimately ensuring that technology strengthens, rather than diminishes, human intellectual capacity.