

AI Teaching Assistants in Online Discussion Forums and Peer-Supported Self-Regulated Learning

Richard Lee Davis^{1,†}, Xiner Liu², Maciej Pankiewicz², Andreas Kissmehl¹,
Olov Engwall¹, Ryan S. Baker³, and Olga Viberg^{1,3,†}

¹ KTH Royal Institute of Technology, Stockholm, Sweden
`{rldavis,kismehl,engwall,oviberg}@kth.se`

² University of Pennsylvania, Philadelphia, PA, USA
`{xiner,mpank}@upenn.edu`

³ Adelaide University, Adelaide, Australia
`ryanshaunbaker@gmail.com`

Abstract. AI-powered teaching assistants (AI TAs) are increasingly deployed in online course forums, yet little is known about how their presence relates to students' self-regulated learning (SRL) strategies or academic outcomes. This study examines four different university courses with different AI TA involvement: a human-only forum and three courses supported by an AI TA, to explore associations between AI TA presence, SRL composition, discussion dynamics, and grade results. Using sentence-level SRL coding across approximately 28,000 student-authored sentences and student-level outcome modeling, we report three main results. First, in the two high-exposure AI courses, the AI TA dominated assignment threads as first responder, and its presence was associated with reduced peer participation; students in the AI-supported courses produced fewer reply posts; and peer contributions to assignment threads declined over the semester. Second, this shift in responder ecology was accompanied by a decline in peer co-regulatory discourse, which was related to lower aggregate levels of regulatory language in AI-supported forums compared with the human-only baseline. Third, in the AI-supported courses, students' self-regulation was the strongest predictor of grade, but peer interaction moderated this relationship: self-regulation was a stronger predictor of grade for students whose threads attracted more peer replies. The overall grade penalty associated with AI TA interaction was concentrated among highly self-regulated students who received less peer co-regulation, suggesting that the displacement of peer dialogue may weaken the support structure through which self-regulation relates to academic outcomes.

Keywords: Self-regulated learning, Co-regulation, Online course forums, AI teaching assistant, STEM higher education

[†]These authors contributed equally to this work and share first authorship.

1 Introduction

Online discussion forums are central in many university courses [15], particularly in technically demanding STEM domains such as computer science (CS) education [6, 16]. In these contexts, students use forums to ask questions, seek clarification, troubleshoot problems, and discuss assignments [12, 6]. Beyond their communicative function, forums also serve as learning spaces, in which discourse can externalize key aspects of self-regulated learning (SRL) [32], including help-seeking, elaboration, organization, effort regulation, and critical thinking. Previous research has used forum interactions to study how learners engage in and regulate their learning in collaboration with peers and instructors [31, 14].

How students participate in forums, and which SRL processes become visible there, is shaped by the structure of interaction [4, 22]: who responds to a question, how quickly support is provided, and whether a discussion expands or closes. Forums can thus be understood as interactional systems in which learning emerges through the interplay of multiple actors and roles. In these settings, peers often contribute co-regulatory support by elaborating, questioning, and explaining, while instructors and teaching assistants (TAs) provide more authoritative forms of scaffolding [25, 3, 33, 10].

Advances in generative AI (GenAI) have led to the increasing integration of AI-mediated support into online discussion environments, including AI agents acting as TAs [20, 16] that offer clear advantages in responsiveness and availability. In contrast to human teachers, AI TAs can respond instantly to multiple student queries and consistently address assignment-related questions. At the same time, they introduce a qualitative shift in how forum interaction unfolds. When students receive immediate answers from an AI TA rather than waiting for peers or instructors, the ecology of participation may change: AI may alter who responds, how long discussions continue, and whether students engage in the kinds of exchanges through which SRL is expressed, practiced, and supported.

Despite the growing adoption of AI TAs [20], their relationship with SRL, participation patterns, and academic outcomes in discussion forums remains poorly understood. This gap is significant because SRL is a key competency for success in online learning and a robust predictor of academic performance [29]. If AI systems change how students seek help, articulate uncertainty, elaborate on problems, or interact with others, they may reshape learning processes in ways that are not fully captured by conventional performance metrics. Existing research has emphasized grades, adoption, or system usage [21], with less attention to how AI support reconfigures the discourse and interactional structures through which regulation and co-regulation of learning occur.

In this paper, we address this gap by conceptualizing AI integration as a reconfiguration of forum ecology rather than a uniform intervention. We treat discussion forums as multi-actor environments in which students, instructors, and AI systems jointly shape interaction patterns and the student learning process. From this point of view, the key question is not whether AI increases or decreases SRL overall, but how it redistributes regulatory activities between actors, thread types, and interaction contexts. This framing allows us to distinguish

between *self-regulation*, which students express in their own learning processes, and *co-regulation*, which emerges when others scaffold those processes through dialogue [9].

Over three levels, we analyzed these issues through four educational technology courses with varying levels of AI TA integration: one human-only forum, one limited-exposure AI condition, and two courses with more extensive AI TA interaction. *First*, we examine how SRL discourse differs across course contexts, testing whether AI-supported forums exhibit a uniform SRL profile or a heterogeneous reconfiguration. *Second*, we analyze how patterns of regulatory language differ between AI-supported and human-only courses and how they develop over time. *Third*, we examine student-level relationships among self-regulation, co-regulation, peer interaction, and academic performance.

We pose the following research questions: **RQ1.** How does the presence of an AI teaching assistant influence participation patterns and peer dialogue in online discussion forums? **RQ2.** How do patterns of self-regulatory and co-regulatory language differ between AI-supported and human-only courses, and how do they develop over the semester? **RQ3.** How are self-regulation, co-regulation, and peer interaction related to academic performance in AI-supported courses?

The **contributions** of this work are threefold. **First**, we show that the AI TA’s presence as a dominant first responder in assignment threads is associated with reduced peer follow-on participation, a pattern that grows over the semester. **Second**, we show that aggregate levels of regulatory language are lower in AI-supported courses, and that this difference is accounted for by the declining share of peer repliers, who carry the majority of co-regulatory discourse. **Third**, we find that students’ own self-regulation is a predictor of academic performance, but that peer interaction moderates this relationship: self-regulation is a stronger predictor of grade when accompanied by peer dialogue. These results highlight the importance of attending to peer co-regulation when designing AI-supported learning environments.

2 Background

2.1 SRL in STEM Education

STEM education, in particular computing education, requires learners to coordinate multiple cognitive and metacognitive processes, including understanding abstract concepts, constructing solutions, debugging iteratively, and evaluating outcomes [18]. Given the diversity in students’ prior knowledge and problem-solving skills, learners must actively regulate their learning processes to succeed [34]. SRL is therefore a critical component of programming education [27, 5].

SRL comprises cognitive and metacognitive strategies such as elaboration, organization, and critical thinking as well as resource management strategies including effort regulation and help-seeking [34]. Prior work in programming education has demonstrated that high-performing students tend to engage more frequently in higher-order strategies (e.g., effort-regulation, critical thinking), whereas lower-performing students rely more on basic strategies such as rehearsal

and executive help-seeking [5]. This highlights that differences in learning outcomes are associated not only with the amount of SRL, but also with the types of strategies students employ in SRL programming practices.

The emergence of commercial large language model (LLM)-based tools, such as Claude and ChatGPT, introduces new forms of student support into programming learning environments. These systems can generate code, explain concepts, and assist with debugging, providing immediate and context-sensitive responses [30, 26]. As such, they offer scalable and personalized support that may complement traditional instruction by creating environments that can, in certain conditions, become conducive to the students' acquisition of programming skills.

However, interactions with LLMs may influence core SRL processes and learners' engagement [11, 13]. Students engage with LLMs in heterogeneous ways, with variation in prompting strategies and reliance patterns [30, 8]. Help-seeking, a central SRL strategy, can become more immediate, iterative, and non-linear in AI-supported contexts [28]. Although this may lower barriers to accessing support, it can also enable learners to bypass critical evaluative and reflective steps, as the ease of obtaining solutions may shift effort regulation, potentially reducing deeper cognitive engagement with programming tasks [30].

The increasing integration of LLMs in computing education is transforming not only learning outcomes but also the ways students engage in SRL. As a result, leveraging these tools effectively for learning may necessitate the adoption of new regulatory strategies. Understanding these transformations requires viewing SRL as emerging within sociotechnical learning settings, where students, instructors, and AI systems collectively shape regulatory processes and learning experiences.

2.2 SRL in Online Discussion Forums

Online discussion forums are important spaces where learners actively engage in SRL [32], by externalizing key regulatory processes through discourse, in the form of help-seeking, explanation, and reflection [7, 32]. Help-seeking is evident when students pose questions, request clarification, or solicit assistance from peers and instructors. Explanation emerges as learners articulate their reasoning, elaborate on problem-solving steps, or provide guidance to others. Reflection is demonstrated in students' evaluation of their understanding, reconsideration of prior approaches, and discussion of alternative strategies. Together, these activities make forums valuable sites for observing how learners regulate their cognition and learning processes in interaction with others.

Prior research has used forum data using discourse and learning analytics approaches to better understand these processes [4, 32]. This work treats forum interactions as trace data that capture learners' cognitive and metacognitive activities as they unfold. In particular, studies have applied SRL coding schemes to textual data, enabling the identification and quantification of learning strategies at different levels of granularity, including fine-grained, sentence-level annotation [32]. This moves beyond simple participation metrics and instead focuses on the quality and nature of learners' contributions. Building on this, we posit that online discussion forums constitute windows into students' SRL processes.

The content and structure of forum discourse provide observable evidence of how learners plan, monitor, and regulate their learning with and without AI TAs.

2.3 Teaching Presence in Online Discussion Forums

Interaction in online discussion forums is shaped not only by what is said, but by *who* responds. Prior research has highlighted the importance of teaching presence and support roles in structuring learning processes, particularly in technology-mediated environments [25, 3]. Within this literature, different actors such as peers and instructors, including teachers and TAs, play distinct roles in scaffolding student learning [33, 10].

Instructor responses are typically more authoritative and directive compared to those from peers. As part of teaching presence, instructors guide discourse, provide clarification, and scaffold student understanding by structuring problems and highlighting key concepts [23]. This form of support can help regulate learning by modeling effective strategies and ensuring conceptual correctness. However, instructor intervention may also shape participation patterns, for example by closing down discussion once a definitive answer is provided.

Recently, AI-mediated instructors in the form of AI TAs have emerged as a new form of support actor in online forums [16, 20]. These systems differ from both peers and human instructors in important ways. AI TAs can provide immediate, scalable, and consistent responses to student queries, particularly in assignment-related contexts. From a scaffolding perspective, they have the potential to support help-seeking by lowering barriers to accessing assistance and enabling iterative questioning. At the same time, their responsiveness and perceived authority may alter how students engage in self-regulation in their learning activities, e.g., by altering their help-seeking behavior and interaction, potentially reducing opportunities for learning. In sum, we conceptualize students, instructors (human TAs and course teachers), and AI TAs as a “responder ecology” shaping student SRL in online discussion forums. Rather than viewing support as uniform, this view stresses how differences in responsiveness, authority, and interaction style influence participation and SRL in online forums.

qualitativ

3 Methodology

3.1 Implemented AI TA

JeepyTA is a platform extending educational discussion forums with functionality enabled by LLMs. Developed at the Penn Center for Learning Analytics at the University of Pennsylvania in 2023, the platform has since been used by faculty at five universities across 29 courses internationally [1].

The central feature of JeepyTA is an AI TA that participates in forum discussions by responding to student posts. Each response is generated through a request to an LLM that contains two primary components: (1) instructions specifying how the response should be generated and (2) contextual information. The instructions are implemented as category-specific prompts that define

the intended goals, rules, and constraints of the AI TA’s responses. Also, instructors can configure visibility settings, including whether AI TA’s responses are published immediately or require instructor approval before becoming visible.

The contextual information always includes the student’s post and may additionally include other materials such as prior discussion history, instructors’ posts, the contents of files attached to the post, and materials attached to the category as a knowledge base. Context construction is supported by a retrieval-augmented generation (RAG) approach, in which textual materials are converted into embeddings to enable semantic search and retrieval of content relevant to the current discussion.

For the courses analyzed in this study, JeepyTA relied on OpenAI GPT models available at the time of deployment: GPT-3.5 Turbo for DS AI, GPT-4 in BD AI, and GPT-4 Turbo in AL AI.

3.2 Datasets

The dataset comprises 28,123 sentence-level units aggregated across four graduate-level courses within data science and educational design, with one human-only condition (DS Human; a data science course) and three courses including an AI TA (respectively, DS AI, the same DS course with AI TA incorporated; BD AI, a course on big data and education; and AL AI, an educational design course focused on adaptive learning systems). Each forum post included a course identifier, thread name, anonymized user ID, and time of posting. The posts were first split into individual sentences by applying the NLTK punkt model [2] and cleaned from HTML artifacts with BeautifulSoup¹. Following a qualitative AI-supported coding approach [17], we developed and applied a codebook (<https://shorturl.at/bsdSa>) in a multi-stage validation procedure. We established a ground truth using a random sample of 100 instances, which were independently annotated by two human coders, resulting in satisfactory inter-rater reliability. (Cohen’s $\kappa > 0.85$, all codes) while ensuring a minimum prevalence threshold of 5% in each label. A few-shot prompt was iteratively crafted for an LLM (GPT-4o, via API) to code the ground truth until agreement with the human-established ground truth exceeded $\kappa = 0.70$ across all constructs. Subsequently, we applied the codes resulting from LLM to the full data set. We further gathered normalized course grade data for each student as a percentage ratio of points obtained relative to the maximum score on the course assignments.

4 Results

The four courses examined exhibit heterogeneous patterns of student- and AI-TA interactions (Table 1). The DS Human course serves as the human-only baseline; the DS AI TA course represents a low-penetration AI context, with only 20% of students interacting with the AI TA; while the BD AI and AL AI courses show high levels of interactions between students and AI TA, reaching 89% and

¹ <https://www.crummy.com/software/BeautifulSoup/>

95% of students, respectively. The remaining analyses therefore compare three conditions: DS Human as the human-only baseline, and BD AI and AL AI as the two AI-supported courses.

Table 1. Study design: four courses with varying AI TA integration. Dashes indicate metrics not applicable to the human-only course.

Course	Condition	Students	AI TA posts	Students reached	Posts followed	Student sentences	Percent SRL sentences
DS Human	Human-only	40	—	—	—	6,491	52%
DS AI	Human+AI (limited)	56	23	20%	1.5%	8,086	60%
BD AI	Human+AI	45	248	89%	41%	4,756	51%
AL AI	Human+AI	39	177	95%	21%	8,687	55%

4.1 How Does the AI TA Change Forum Dynamics?

By design, the AI TA was intended to respond in assignment threads, and these were hence more likely to receive a first response from the AI TA, $\chi^2(1) = 178.8$, $p < .001$, OR = 62.7, 95% CI [27.3, 169.0]. In both BD AI TA and AL AI TA, the AI TA was the first responder in 97% of student-started assignment threads (Figure 1, left). Also by design, the AI TA was much faster than human responders (Figure 1, right). The median time to first response in student-started assignment threads was 0.92 minutes for AI TA-first threads ($n = 195$; IQR [0.67, 1.21]), compared with 490 min for peer-first threads in the human TA course ($n = 75$; IQR [117, 942]) and 364 min for teacher-first threads in the human TA course ($n = 38$; IQR [89.8, 1068]), $\chi^2(2) = 209.62$, $p < .001$.

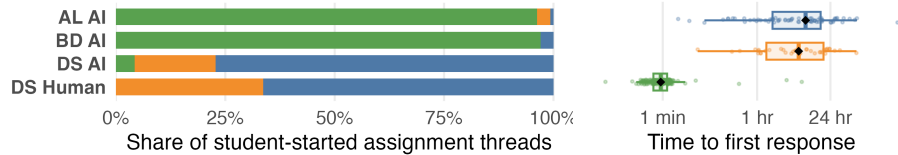


Fig. 1. (Left) first responder in student-started assignment threads and (Right) time to first response, in log-scale, by **students**, **teachers**, and **AI TA**.

Next, we turn to the impact on students' forum activity. We firstly observe that students in the AI TA courses posted fewer times on average than students in the human-only course. Students in the human-only course averaged 23.05 posts each, whereas students in the two AI-supported courses averaged 15.85 posts each, $\chi^2(1) = 19.08$, $p < .001$. This difference did not come from thread starts, which were similar across conditions (means = 4.93 vs. 4.88; $p = .73$), but from follow-on replies: students in the human-only course produced more reply posts (mean = 18.2, median = 18) than students in the AI-supported courses (mean = 10.9, median = 10.5), $\chi^2(1) = 26.31$, $p < .001$ (Figure 2).

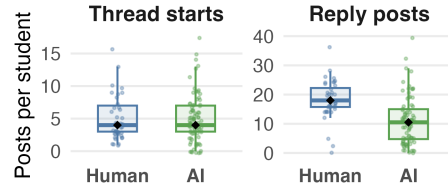


Fig. 2. Posts per student in the human-only course versus the two AI-supported courses, decomposed into thread starts and reply posts.

4.2 Relationship between Interaction Partners and Grades

We next investigated how students' conversation partners were related to grade outcomes in the two AI-supported courses. For each student, we identified the threads that a student started and then calculated three interaction-partner measures: the proportion of responses in those threads that came from a teacher, the proportion that came from a peer, and the proportion that came from the AI TA. These measures therefore describe who populated the conversations that students themselves initiated. We then fit three separate linear regression models, each predicting course-standardized grade from one interaction-partner measure while controlling for posting volume and course. For the linear models reported here and below, we use HC3 robust standard errors [19]; the logistic models later in the section use model-based standard errors.

We found that students who started threads containing a higher proportion of teacher responses earned higher grades ($\beta \approx 2.12$, $p = .009$), whereas students who started threads that then contained a higher proportion of AI TA responses earned lower grades ($\beta \approx -1.38$, $p = .038$). Peer response proportion showed no clear relationship with grade ($\beta \approx 0.02$, $p = .99$) (Figure 3, left).

Importantly, this pattern is driven by who responds in a student's threads, not by where the student posts. A separate model predicting grade from the proportion of a student's activity occurring in assignment threads showed no significant relationship ($\beta \approx -0.69$, $p = .282$), even though assignment threads are precisely where the AI TA is active.

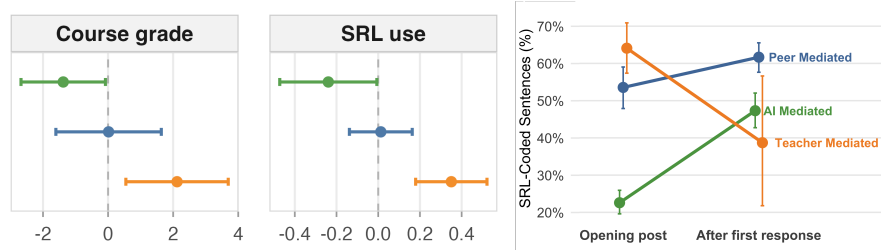


Fig. 3. Estimated association between first responder in assignment threads (AI TA, student, teachers) and (Left) Course grade or (Middle) SRL use. (Right) Proportion of sentences coded as SRL in the original poster's opening post versus subsequent posts, by thread type. Error bars show 95% bootstrap confidence intervals.

Regulatory Language in AI-Supported vs. Human-Only Courses

The SRL strategies coding categories used in this study (help-seeking, elabo-

ration, critical thinking, organization, and effort regulation) capture regulatory processes that may serve different functions depending on context. When a student uses regulatory language in a thread that they started (e.g., articulating a plan, monitoring their understanding, or seeking targeted help), these processes constitute self-regulation. When the same language appears in a student’s reply to another student’s thread, it is consistent with co-regulation: one learner scaffolding another’s regulatory activity [9]. This distinction matters because one of the AI TA’s effects, as shown in section 4.1, was to displace peer replies, and peer replies are the mechanism through which co-regulation occurs. We examine: 1) how aggregate levels of regulatory language differ between conditions, 2) the student-level relationship between regulatory language and grades.

Lower SRL in AI-Supported Courses. We first asked whether the AI-supported courses look different from the human-only baseline on overall levels of regulatory language. A logistic regression predicting whether each sentence was coded as SRL, with condition (human-only vs. AI-supported) as the predictor, showed that the odds of a sentence being coded as SRL were significantly higher in the human-only course than in the two pooled AI-supported courses ($OR = 0.65$, 95% CI [0.62, 0.70], $p < .001$). This difference was concentrated in assignment-thread opening posts. In AI-mediated threads, the original poster’s opening move contained less SRL than comparable openings in peer-mediated or teacher-mediated threads. After the opening post, SRL rose in AI-mediated threads but fell in teacher-mediated threads (Figure 3, right).

There were also significant shifts in SRL over the course of the semester. We split each student’s coded sentences into the first (H1) and second halves (H2) of the semester and computed the proportion of sentences coded as SRL in each half. In the AI-supported courses, the proportion of sentences coded as SRL declined from the first half of the semester (mean = 49%) to the second half (mean = 36%), $t(67) = -6.03$, $p < .001$. In the human-only course, SRL trended slightly upward instead (54% to 61%; $t(38) = 2.02$, $p = .050$). This decline persisted even when excluding the assignment-thread opening posts (mean decrease = -19 percentage points, $t(61) = -6.00$, $p < .001$) (Figure 4).

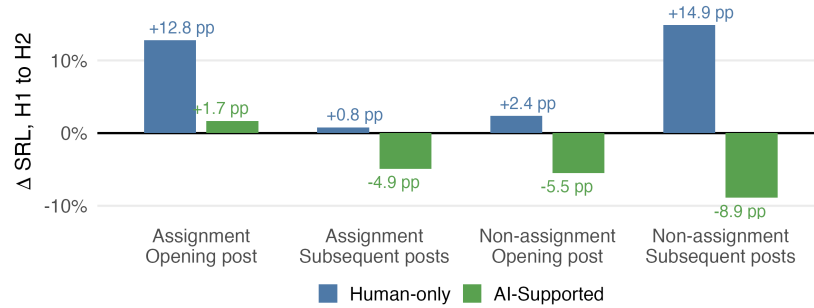


Fig. 4. Change in the proportion of sentences coded as SRL from the first half to the second half of the semester, decomposed by thread type and post position.

The SRL Decline Reflects Lost Peer Co-regulation. To understand why SRL declined, we examined follow-up sentences in assignment threads.

In the pooled AI-supported courses, peer repliers produced approximately 70% SRL-coded sentences (70.2% in H1; 68.5% in H2), whereas thread starters produced much less (26.9% in H1; 20% in H2). As the SRL-coded discourse appears in peer replies, it is more consistent with co-regulation rather than self-regulation [9].

What changed over the semester was not the rate at which each role produced SRL, but how much each role contributed to the conversation. The proportion of follow-up sentences written by peer repliers fell from 91% in the first half to 45% in the second half. Because SRL rates within each role were relatively stable across halves (peer repliers: 70.2% to 68.5%; thread starters: 26.9% to 20%), the decline in overall regulatory language was driven by the shrinking share of peer repliers rather than by changes in SRL rates within each role.

4.3 Self-Regulation, Co-regulation, and Grades

The previous section established that AI-supported courses exhibit lower aggregate levels of regulatory language, and that this decline is driven by the displacement of peer repliers, who are the primary carriers of co-regulatory discourse. We now turn to the student level and ask how these regulatory processes relate to individual course outcomes.

Separating Self-Regulation from Co-regulation. We decomposed each student’s regulatory language into three components based on where it occurs. Self-regulation captures a student’s own regulatory language in threads they started. Received co-regulation captures the regulatory language that human participants (excluding the AI TA) contributed to those same threads. Provided co-regulation captures a student’s co-regulatory language when replying in other students’ threads. Each measure is a proportion of sentences coded as SRL within its respective context. For each measure, we fit a separate linear regression predicting course-standardized grade from that measure while controlling for posting volume and course, with HC3 robust standard errors (Figure 5).

Self-regulation was a significant positive predictor of grade ($\beta = 2.63$, 95% CI [0.49, 4.76], $p = .017$, $n = 66$). Neither received co-regulation ($\beta = 0.66$, 95% CI [-1.21, 2.54], $p = .482$, $n = 64$) nor provided co-regulation ($\beta = 0.43$, 95% CI [-1.52, 2.38], $p = .660$, $n = 67$) significantly predicted grade.

Peer Interaction Moderates Self-Regulation vs. Grade. Self-regulation predicts grade, but neither receiving nor providing co-regulation does so independently. However, one of the AI TA’s effects was to displace peer replies. If peer dialogue sustains or amplifies the benefit of self-regulation, self-regulation should predict grade more strongly for students who also received more peer replies. We tested this by fitting a linear regression predicting course-standardized grade from self-regulation share, log peer reply volume in the student’s started threads, their interaction, and course, with HC3 robust standard errors ($n = 66$). The interaction term was significant ($\beta = 1.45$, 95% CI [0.31, 2.60], $p = .013$). The interaction model explained more variance than the main-effects-only model ($R^2 = 0.17$ vs. 0.10; F -test $p = .029$ for the ΔR^2). Self-regulation was a stronger predictor of grade for students whose threads contained more peer replies.

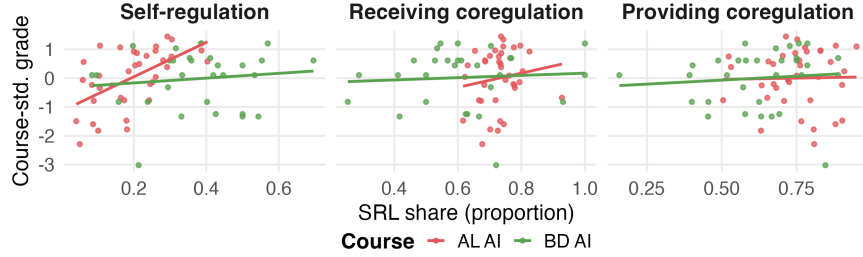


Fig. 5. Relationship between the proportion of SRL-coded sentences and course-standardized grade for self-regulation (left), received co-regulation (center) and provided co-regulation (right) .

Table 2. Mean course-standardized grade by early self-regulation and received co-regulation (median split) in the pooled AI-supported courses ($n = 62$).

	High received co-regulation	Low received co-regulation	Δ
High early self-reg ($n = 30$)	+0.90 ($n = 9$)	+0.16 ($n = 21$)	+0.74 ($p = .004$)
Low early self-reg ($n = 32$)	-0.19 ($n = 22$)	-0.55 ($n = 10$)	+0.36 ($p = .369$)

Who Is Impacted by Lost Co-regulation? The moderation analysis showed that peer interaction amplifies the benefit of self-regulation. Given that the AI TA displaces peer replies, we next asked which students might be impacted by the potential loss of peer co-regulation. We split students at the median on two measures: early self-regulation (the proportion of SRL-coded sentences in students’ own threads during H1) and received co-regulation (the SRL share of human participants’ replies in their started threads, excluding AI TA). We then compared mean course-standardized grades within each row of the resulting 2×2 table using independent-samples t -tests ($n = 62$; Table 2).

Among students with high early self-regulation, those who received above-median co-regulation earned significantly higher grades than students who received below-median co-regulation ($\Delta = 0.74$ SD, $t(26) = 3.15$, $p = .004$). Among students with low early self-regulation, the same contrast was not significant ($\Delta = 0.36$ SD, $t(19) = 0.92$, $p = .369$). Received co-regulation was thus more consequential for students who entered the course with stronger self-regulatory habits. This finding also connects back to the overall AI TA grade penalty reported in Section 4.2. Removing the 17 high-early-self-regulation students whose started threads had above-median AI TA reply share and re-estimating the linear regression from Section 4.2, the AI TA coefficient dropped from $\beta \approx -1.38$ ($p = .038$) to $\beta \approx -1.29$ ($p = .454$). The grade penalty associated with the AI TA is thus concentrated among the students whose outcomes are most sensitive to co-regulatory support.

5 Discussion and Conclusion

Across three levels of analysis, the results point to a consistent pattern: the AI TA’s presence was associated with reduced peer dialogue and lower levels of co-regulatory discourse, and peer interaction moderated the relationship between self-regulation and grades. We summarize these findings in terms of the research questions, then turn to interpretation and implications.

RQ1: How does the presence of an AI teaching assistant influence participation patterns and peer dialogue in online discussion forums? Students in the AI-supported courses produced fewer reply posts than in the human-only course, while their thread-starting rates were comparable. The share of follow-up sentences contributed by peer repliers fell sharply over the semester in the AI-supported courses but remained stable in the human-only course.

RQ2: How do patterns of self-regulatory and co-regulatory language differ between AI-supported and human-only courses, and how do they develop over the semester? Aggregate levels of human regulatory language were lower in the AI-supported courses and declined over the semester, while remaining stable in the human-only course. The decline was driven by the shrinking share of peer repliers. As peer-authored regulatory language appeared in replies to other students’ threads, this is consistent with decreased co-regulation [9], which emerged as important to understanding observed patterns.

RQ3: How are self-regulation, co-regulation, and peer interaction related to academic performance in AI-supported courses? A higher proportion of AI TA responses in students’ threads was associated with lower grades. Self-regulation was the strongest predictor of grade, but peer interaction moderated this relationship: self-regulation predicted grades more strongly for students whose threads attracted more peer replies. The students most affected by reduced peer dialogue were those with high early self-regulation; among this group, those who received more co-regulation earned substantially higher grades. This difference was absent among students with low early self-regulation. The overall AI TA grade penalty became non-significant after removing these high self-regulation students, suggesting that the AI TA’s association with lower grades operates through the displacement of peer co-regulation rather than through a direct relationship with the AI TA itself.

Interpretation. The results describe a coherent set of associations across three levels of analysis. At the forum level, the AI TA’s rapid responses in assignment threads co-occurred with reduced peer follow-on participation. At the discourse level, this reduction in peer participation resulted in lower aggregate levels of regulatory language, since peer repliers were the primary contributors of co-regulatory discourse. At the student level, self-regulation predicted grades more strongly when accompanied by peer interaction, and students with high early self-regulation were most sensitive to co-regulation influence on grades.

Traditional SRL models emphasize regulation as an individual competency [34]. We entered this study with that framing, examining how SRL differed across AI-supported and human-only courses. However, the patterns we observed pointed

to the importance of co-regulatory processes, in which one learner scaffolds another’s regulatory activity [9], as a complement to individual self-regulation. When the AI TA displaced that dialogue, the association between self-regulation and grades weakened. Why this effect was concentrated among high self-regulators rather than low self-regulators is not clear from the data, and unmeasured characteristics (e.g., prior knowledge, motivation) may account for the pattern. We note that the observational design does not allow us to establish whether these relationships are causal.

Design Implications. The AI TA’s speed and consistency are clear operational benefits, but they co-occurred with reduced peer dialogue and lower levels of co-regulatory discourse. If this pattern reflects a genuine effect, it points to a trade-off: the same responsiveness that makes AI TAs attractive may displace the peer interactions through which co-regulation occurs. Designers of AI-supported forums should consider how to preserve opportunities for peer dialogue. One concrete lever is response timing: wait-time research shows that delaying an authoritative reply increases the quantity and complexity of student contributions [24]; manipulating AI TA response latency in asynchronous forums is therefore a natural follow-up experiment.

Limitations and Future Directions. Several limitations should be acknowledged. First, we rely on forum discourse as a proxy for SRL, capturing only regulatory processes that students articulate in writing. Internal regulatory activities remain unobserved, and future work could incorporate multimodal or trace-based data. Second, the focus on graduate-level courses in data science and educational design may limit generalizability; students in these contexts may exhibit relatively advanced self-regulatory skills, and their interaction patterns with AI systems may differ from those of students in other domains or levels. Third, the observational design limits causal inference. The associations we report may reflect course-level or student-level confounds rather than effects of the AI TA itself, and readers should interpret them accordingly.

Experimental or quasi-experimental designs, for example, randomizing AI TA availability within a single course, would allow researchers to test whether the displacement of peer dialogue follows directly from AI intervention. Additionally, the finding that students with high early self-regulation appear most sensitive to co-regulation warrants replication. If this pattern holds, it would challenge the assumption that students with stronger self-regulation are the more self-sufficient learners [34].

In conclusion, these findings suggest that the co-regulatory role of peer contributions deserves attention in the design of AI-supported learning environments, and that the speed and availability of AI TAs should be weighed against the possible reduction in peer interactions that support regulatory processes and ultimately student academic success in higher education.

References

1. Baker, R.S., Liu, X., Shah, M., Pankiewicz, M., Kim, Y.J., Lee, Y., Porter, C.: Generative ai as a teaching assistant. OECD Digital Education Outlook (2026)

2. Bird, S., Klein, E., Loper, E.: Natural language processing with Python: analyzing text with the natural language toolkit. " O'Reilly Media, Inc." (2009)
3. Blaine, A.M.: Interaction and presence in the virtual classroom: An analysis of the perceptions of students and teachers in online and blended advanced placement courses. *Computers & Education* **132**, 31–43 (2019)
4. Chen, B., Poquet, O.: Uncovering socio-temporal dynamics in online discussions: An event-based approach. *Australasian Journal of Educational Technology* **39**(6), 1–16 (2023)
5. Cheng, G., Zou, D., Xie, H., Wang, F.L.: Exploring differences in self-regulated learning strategy use between high-and low-performing students in introductory programming: An analysis of eye-tracking and retrospective think-aloud data from program comprehension. *Computers & Education* **208**, 104948 (2024)
6. Cloude, E.B., Baker, R.S., Fouh, E.: Online help-seeking occurring in multiple computer-mediated conversations affects grades in an introductory programming course. In: LAK23: 13th International Learning Analytics and Knowledge Conference. pp. 378–387 (2023)
7. Du, Z., Wang, F., Wang, S., Xiao, X.: Enhancing learner participation in online discussion forums in massive open online courses: The role of mandatory participation. *Frontiers in psychology* **13**, 819640 (2022)
8. Güner, H., Er, E.: Ai in the classroom: Exploring students' interaction with chatgpt in programming learning. *Education and Information Technologies* **30**(9), 12681–12707 (2025)
9. Hadwin, A.F., Järvelä, S., Miller, M.: Self-regulated, co-regulated, and socially shared regulation of learning. *Handbook of self-regulation of learning and performance* pp. 65–84 (2011)
10. Hmelo-Silver, C.E., Duncan, R.G., Chinn, C.A.: Scaffolding and achievement in problem-based and inquiry learning: a response to kirschner, sweller, and. *Educational psychologist* **42**(2), 99–107 (2007)
11. Huang, X., Zhang, S.: Engagement in llm chatbot-supported learning: The pivotal roles of genai competency and emotion. *Computers and Education: Artificial Intelligence* p. 100559 (2026)
12. Ihantola, P., Vihavainen, A.e.a.: Educational data mining and learning analytics in programming: Literature review and case studies. *Proceedings of the 2015 ITiCSE on working group reports* pp. 41–63 (2015)
13. Kirkeby-Hinrup, A., Stenseke, J.: The psychology of llm interactions: the uncanny valley and other minds. *Journal of Psychology and AI* **1**(1), 2457627 (2025)
14. Kizilcec, R.F., Pérez-Sanagustín, M., Maldonado, J.J.: Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & education* **104**, 18–33 (2017)
15. Leon, L.R., Plowright-Pepper, L.: The emotional labour of online course discussion forums: some autoethnographic insights. *Open Learning: The Journal of Open, Distance and e-Learning* **41**(1), 67–82 (2026)
16. Liu, X., Pankiewicz, M., Gupta, T., Huang, Z., Baker, R.S.: A step towards adaptive online learning: Exploring the role of gpt as virtual teaching assistants in online education. In: *Future of Learning with Large Language Models*, pp. 149–166. CRC Press (2025)
17. Liu, X., Zambrano, A.F., Baker, R.S., Barany, A., Ocumpaugh, J., Zhang, J., Pankiewicz, M., Nasir, N., Wei, Z.: Qualitative coding with gpt-4: Where it works better. *Journal of Learning Analytics* **12**(1), 169–185 (Mar 2025)

18. Loksa, D., Margulieux, L., Becker, B.A., Craig, M., Denny, P., Pettit, R., Prather, J.: Metacognition and self-regulation in programming education: Theories and exemplars of use. *ACM Transactions on Computing Education (TOCE)* **22**(4), 1–31 (2022)
19. Long, J.S., Ervin, L.H.: Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician* **54**(3), 217–224 (2000)
20. Pozdniakov, S., Brazil, J.e.a.: From misunderstandings to learning opportunities: Leveraging generative ai in discussion forums to support student learning. In: *International Conference on Artificial Intelligence in Education*. pp. 291–298. Springer (2025)
21. Qiao, Y., Shihab, M.I.H., Hundhausen, C.: A systematic literature review of the use of genai assistants for code comprehension: implications for computing education research and practice. *ACM Transactions on Computing Education* **26**(2), 1–33 (2026)
22. Rabbany, R., ElAtia, S., Takaffoli, M., Zaïane, O.R.: Collaborative learning of students in online discussion forums: A social network analysis perspective. In: *Educational data mining: Applications and trends*, pp. 441–466. Springer (2013)
23. Richardson, J.C.e.a.: Conceptualizing and investigating instructor presence in online learning environments. *International Review of Research in Open and Distributed Learning* **16**(3), 256–297 (2015)
24. Rowe, M.B.: Wait time: Slowing down may be a way of speeding up! *Journal of Teacher Education* **37**(1), 43–50 (1986)
25. Shea, P., Hayes, S., Vickers, J.: Online instructional effort measured through the lens of teaching presence in the community of inquiry framework: A re-examination of measures and approach. *International Review of Research in Open and Distributed Learning* **11**(3), 127–154 (2010)
26. Shen, J.H., Tamkin, A.: How ai impacts skill formation. *arXiv:2601.20245 [cs.LG]* (2026)
27. Silva, L., Mendes, A., Gomes, A., Fortes, G.: What learning strategies are used by programming students? a qualitative study grounded on the self-regulation of learning theory. *ACM Transactions on Computing Education* **24**(1), 1–26 (2024)
28. Viberg, O., Feldman-Maggor, Y., Wong, J.: Efficiency vs. effectiveness: Self-regulated learning with llm-mediated help-seeking. *Learning Letters* **8**, 60–60 (2026)
29. Viberg, O., Khalil, M., Baars, M.: Self-regulated learning and learning analytics in online learning environments: A review of empirical research. In: *Proceedings of the 10th International conference on LAK*. pp. 524–533 (2020)
30. Viberg, O., Wong, J.e.a.: Chatting with code: Exploring llms as learning partners in programming education. In: *International Conference on Artificial Intelligence in Education*. pp. 453–461. Springer (2025)
31. Wang, Z., Chen, X.e.a.: How does the online discussion intervention influence learners’ engagement, emotions, and motivation in a blended learning environment? learners’ need for cognition matters. *BMC psychology* **13**(1), 1122 (2025)
32. Wise, A.F., Hsiao, Y.T.: Self-regulation in online discussions: Aligning data streams to investigate relationships between speaking, listening, and task conditions. *Computers in Human Behavior* **96**, 273–284 (2019)
33. Yu, H., Chen, Y., bin Ismail, I.M.: From scaffolding to success: How instructor pedagogical support and collaborative classroom interaction drive scholastic motivation in programming education. *Acta Psychologica* **259**, 105289 (2025)
34. Zimmerman, B.J.: Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice* **41**(2), 64–70 (2002)