Metacognition and Learning (2026) 21:24
https://doi.org/10.1007/s11409-026-09468-y

How rubric performance-level order and feedback shape students' real-time self-regulation and self-efficacy

Ernesto Panadero[1,2,3] · **Leire Pinedo**[2] · **Alazne Fernández-Ortube**[2] · **David Zamorano**[2] · **Lucía Barrenetxea-Minguez**[2] · **Iván Sánchez-Iglesias**[4]

Received: 5 July 2025 / Accepted: 2 April 2026
© The Author(s) 2026

Abstract
This study examines how the order of rubric performance levels (high-to-low vs. low-to-high) and four types of feedback (no feedback, rubric-based, product, and process) influence university students' self-regulated learning (SRL) and self-efficacy during an academic task. A total of 120 first-year students completed two landscape analysis tasks supported by rubrics while verbalizing their thoughts through think-aloud protocols. Self-efficacy was measured before and after the experimental session. Students predominantly produced positive self-regulatory actions across the five task phases, displaying a W-shaped pattern that reflected shifts in task demands: higher regulatory engagement during rubric reading and written analysis, and lower engagement during oral description phases. Negative actions were less frequent and concentrated in the initial rubric reading and first written analysis. Neither performance-level order nor feedback type significantly affected positive or negative SRL actions; the only significant within-subjects factor was task phase. Self-efficacy decreased across all conditions, suggesting a recalibration of confidence in response to task demands and rubric criteria. These findings indicate that moment-to-moment SRL engagement is primarily shaped by the nature of the task phase rather than by the specific scaffolding configuration, and underscore the importance of process-oriented methods for capturing dynamic regulatory patterns in rubric-supported learning environments.

Keywords Rubric · Feedback · Self-regulated learning · Think-aloud protocols · Self-efficacy

✉ Leire Pinedo
leire.research@gmail.com; leire.pinedo@opendeusto.es

[1] Centre for Assessment Research Policy and Practice in Education (CARPE), School of Policy and Practice, Institute of Education, St. Patrick's Campus, Dublin City University, Dublin, Ireland

[2] Education Regulated Learning and Assessment group (ERLA group), Facultad de Educación y Deporte, Universidad de Deusto, Bilbao 48007, Spain

[3] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

[4] Department of Psychobiology and Behavioral Sciences Methods, School of Psychology, Complutense University of Madrid, Madrid, Spain

Published online: 30 May 2026

Supporting learners in becoming effective self-regulators is a longstanding goal of educational research and practice (Greene, 2018). As contemporary classrooms increasingly emphasize autonomy and higher-order thinking, there is a growing need for instructional approaches that help students to monitor, evaluate, and adapt their learning strategies. Among the most widespread assessment-based supports are rubrics and formative feedback, which provide external scaffolds intended to clarify expectations, guide self-assessment, and foster self-regulated learning (Brookhart, 2018; Panadero et al., 2023). Despite the widespread adoption of both, rubrics and feedback, their actual effects on students' self-regulated learning (SRL) processes during task engagement remain only partially understood.

Recent studies have highlighted that rubrics and feedback can have a dual role: while they may enhance students' cognitive and motivational engagement, they can also generate confusion or lead to superficial compliance if learners are uncertain about how to interpret or use them (Andrade, 2005; Dawson, 2017). Furthermore, little is known about how rubric design features (e.g., order of performance levels presented in a rubric) or variations in feedback type influence the real-time deployment of self-regulation (Panadero et al., 2025a). Addressing these questions requires methods that can document what students actually do while working with a rubric and how their perceptions change as the task progresses. To meet this need, the present study integrates think-aloud protocols with self-report measures and examines how rubric configuration and feedback shape learners' SRL and self-efficacy.

Self-regulated learning in academic contexts

Self-regulated learning (SRL) refers to the self-generated thoughts, feelings and actions that students use to manage their learning processes in a goal-directed and cyclical manner (Zimmerman, 2000). An essential aspect of SRL is its dynamic and context-sensitive nature. Rather than functioning as a fixed trait, SRL unfolds in cycles shaped by the characteristics of tasks and learning environments (Winne & Hadwin, 1998). For example, learners may engage in more intensive monitoring when facing complex writing tasks, or they may rely on procedural scripts in familiar problem-solving contexts (Efklides, 2011; De Bruin & van Merriënboer, 2017). This adaptability underscores the importance of studying SRL not only through questionnaires, but also through process-oriented methods capable of revealing strategy use in real time (Winne, 2010).

Research has shown that students differ not only in the frequency with which they regulate their learning, but also in the quality of their regulatory responses. Positive SRL actions involve strategic engagement, constructive motivational self-talk and persistence when facing difficulty, reflecting regulatory processes that support progress and learning. Negative SRL actions, in contrast, involve avoidance, expressions of stress or frustration, and task-minimising strategies that prioritise short-term relief over effective engagement. This distinction parallels the adaptive–maladaptive distinction commonly used in SRL research (Boekaerts & Niemivirta, 2000; Ben-Eliyahu & Bernacki, 2015), but is grounded in empirical work on students' self-messages and regulatory habits (Alonso-Tapia et al., 2014). Thinking-aloud data are particularly suited to capture these positive and negative responses as they unfold in real time, allowing us to identify how learners react to difficulty, interpret feedback and adjust their approach to the task.

Measuring self-regulated learning: combining think-aloud protocols and self-report instruments

Capturing the complexity of self-regulated learning requires methodological approaches that go beyond static self-perceptions. Traditionally, self-report questionnaires have been the predominant tool for measuring SRL, offering practical insights into learners' beliefs about their regulatory tendencies (Boekaerts & Corno, 2005; Zimmerman & Schunk, 2011). While these measures have demonstrated predictive validity (Richardson et al., 2012), they are inherently limited in documenting how regulation unfolds dynamically during task engagement (Winne, 2010).

Process-oriented methods, such as think-aloud protocols, have emerged as essential complements to self-reports. By prompting learners to verbalize their thoughts in real time, think-alouds provide a direct window into cognitive and metacognitive processes (Greene et al., 2011; Azevedo et al., 2018). This approach has been instrumental in uncovering both positive strategies (e.g., planning, monitoring) and negative responses (e.g., avoidance, disengagement) that might remain inaccessible through retrospective questionnaires.

However, each method carries unique strengths and limitations. Self-reports are well-suited to capture more stable, trait-like perceptions of SRL, but can be susceptible to biases such as social desirability or limited metacognitive awareness (Winne, 2020). Conversely, think-alouds provide richer situational detail but may inadvertently influence learners' engagement or introduce reactivity effects (Fox et al., 2011). Combining these complementary data sources can therefore strengthen the interpretive validity of findings, offering a multidimensional perspective on regulation as both a disposition and a situated process (McCardle & Hadwin, 2015).

In the present study, we draw on this multi-method tradition by integrating think-aloud protocols to observe self-regulatory actions during task execution and self-report measures to assess broader motivational and regulatory orientations. This design allows us to explore not only how rubric design and feedback shape real-time SRL behaviors, but also how their self-efficacy is influenced by these instructional supports.

Rubrics as scaffolds for self-regulated learning

Rubrics have gained prominence as versatile tools in educational assessment and instructional design (Dawson, 2017; Panadero et al., 2023). Typically presented as a matrix that articulates performance levels across multiple criteria, rubrics provide learners with explicit expectations and standards of quality (Brookhart, 2018; Dawson, 2017). Rubrics can support SRL by prompting students to reflect on task demands, monitor their progress, and plan improvements, and prior research has shown that they enhance self-regulatory strategies such as goal setting and self-evaluation while also strengthening self-efficacy (Andrade, 2005; Panadero et al., 2023).

Despite these benefits, researchers have noted that rubrics are not universally effective. Their impact depends on both contextual factors (e.g., prior experience, instructional support) and design features, including the clarity of descriptors and the sequence in which performance levels are presented (Jonsson et al., 2025; Panadero et al., 2023). For example, recent evidence suggests that presenting the most proficient performance level first can

shape students' expectations and attentional focus differently compared to starting from the lowest level (Panadero et al., 2025a). However, much of previous research has primarily examined outcomes such as rubric reading patterns or summative performance scores, leaving important questions about real-time self-regulatory processes underexplored. Furthermore, while rubrics are often used in combination with formative feedback, relatively few studies have systematically investigated how different types of feedback interact with rubric use to shape students' self-regulation (for an exception see Panadero et al., 2012). This is a notable gap, considering that in authentic classroom practice, rubrics and teacher feedback are frequently integrated to scaffold learning (Brookhart, 2018; Dawson, 2017). The dynamic interplay between these supports and how they influence both positive and negative SRL remains insufficiently understood.

The role of feedback in supporting self-regulatory processes

Feedback is widely recognized as one of the most powerful drivers of learning, with meta-analytic evidence demonstrating its consistent positive effects on achievement across diverse contexts and disciplines (Hattie & Timperley, 2007). From a self-regulated learning perspective, feedback serves as an external scaffold that informs learners about their current progress relative to goals and standards, helping them to adjust strategies, sustain motivation, and build self-efficacy (Nicol & Macfarlane-Dick, 2006).

The effectiveness of feedback, however, depends on multiple characteristics, including its timing, specificity, focus (process vs. product), and perceived credibility (Shute, 2008). Process feedback, which focuses on strategies and ways to improve rather than on outcomes, is considered more effective for supporting learners' engagement and adaptive regulation, whereas product or purely evaluative feedback can lead to superficial compliance or defensive responses (Hattie & Timperley, 2007).

Importantly, rubrics and feedback are often implemented concurrently in educational settings. Rubrics provide transparent performance criteria, while formative feedback can help students interpret these criteria, identify gaps in their work, and plan next steps (Brookhart, 2018). Yet, research exploring how different combinations of rubric design and feedback types affect real-time regulation and self-perceptions remains limited. As Andrade (2005) noted, without explicit guidance, students may struggle to use rubrics effectively, and feedback can serve as a crucial bridge between awareness of criteria and enactment of improvement strategies.

Researchers have been investigating this interplay. For example, Panadero et al. (2012) showed that orally delivered process feedback combined with rubrics can enhance SRL more effectively than product feedback or rubrics alone. Nevertheless, few investigations have examined how these effects manifest dynamically during the completion of authentic tasks. Capturing such moment-to-moment processes requires observational and process-oriented methods capable of documenting how feedback interacts with other instructional supports to shape self-regulation.

In the present study, we extend this line of inquiry by comparing several feedback types -i.e., product, process, rubric-based, and control-delivered alongside rubrics with differing performance level orders. This approach allows us to explore whether and how feedback type influences positive and negative SRL and self-efficacy, contributing to a more nuanced understanding of feedback as a mechanism for metacognitive scaffolding.

The role of rubrics and feedback in shaping self-efficacy

Self-efficacy refers to the belief in one's capability to perform at a particular level on a task (Bandura, 1997). It is conceptualized as a context-specific belief that is influenced by one's perceived performance during the activity shaped by previous experiences of success or failure and the characteristics of the learning task (Bandura, 1997). Multiple studies using repeated measures during learning activities have shown that self-efficacy is not static, but fluctuates as individuals progress through a task (e.g., Bernacki et al., 2015; Nuutila et al., 2021). For example, in a longitudinal study, Pinedo et al. (2026) found that self-efficacy showed modest changes across occasions, while also responding dynamically to performance outcomes: it increased after successful performance and decreased when students discovered lower-than-expected instructor's scores. These dynamic shifts highlight the importance of considering occasion as a key factor when examining self-efficacy in authentic academic contexts.

Instructional scaffolds such as rubrics and feedback are particularly relevant in shaping self-efficacy judgments because they provide learners with external information about task standards and their proximity to those standards. Rubrics can serve as clear reference frameworks that reduce ambiguity, enabling students to better understand expectations and monitor their progress (Brookhart, 2018). While some studies report that rubric use can increase students' self-efficacy (Panadero et al., 2012), others have found non-significant effects (Andrade et al., 2009). A recent meta-analysis by Panadero and colleagues (2023) found a small but positive overall effect of rubrics on self-efficacy, suggesting that the impact of rubrics may depend on contextual and design-related factors, such as how performance levels are described or sequenced.

Feedback type also plays a critical role in shaping self-efficacy, as feedback that emphasizes progress and mastery can enhance both motivation and confidence (Schunk et al., 2014). In particular, process feedback, which focuses on strategies and understanding rather than outcomes, has been found to increase self-efficacy more than either product feedback or rubric-based feedback (Panadero et al., 2012), which emphasize the final outcome and the degree to which previously established standards have been met, respectively. By directing attention to the underlying strategies instead of the final result, it may help students interpret their performance more constructively and strengthen their belief in their capacity to improve. Despite the importance of feedback for self-efficacy, there is limited evidence directly comparing process, product, and rubric-based approaches. Examining how these different instructional supports influence students' confidence and understanding of the task is therefore crucial for informing practices that effectively promote learning and motivation.

Aim, research questions and hypotheses

The primary aim of this study was to examine how two scaffolding strategies (i.e., the order in which rubric performance levels are presented and the type of feedback students receive) support or hinder students' SRL and self-efficacy during an academic writing activity. This study is part of a broader research programme investigating how rubric-based scaffolding shapes students' engagement with complex academic tasks. Whereas previous publications from the project have focused on visual attention patterns and students' perceptions of rubric design (Panadero et al., 2025a, b), the present work addresses a different set of questions. Specifically, we examine how rubric order and feedback type influence moment-to-moment

self-regulation, as captured through think-aloud protocols, as well as students' self-efficacy. Our research questions (RQ) and hypotheses (H) in this study were:

RQ1 Which positive and negative self-regulatory actions do students verbalize through think-aloud protocols while working with the rubric?

H1. This question is descriptive in nature, aiming to provide a detailed taxonomy of students' SRL.

RQ2 What are the effects of rubric performance level order and feedback type on the frequency of positive and negative self-regulatory actions as captured by think-aloud protocols?

H2. Drawing on prior evidence that top-down rubric structures (presenting the highest performance level first) guide students' attention toward higher standards and support more effective task engagement, we expect that this rubric order will be associated with a higher proportion of positive SRL actions relative to negative ones (elicited via think-aloud). In parallel, we anticipate that process feedback -which provides guidance on how to improve rather than only what was done well or poorly- will further promote positive regulatory responses. Thus, both scaffolding features are expected to influence the balance of positive and negative SRL actions, with the top-down rubric order and process feedback eliciting more productive regulatory patterns.

RQ3 What are the effects of rubric performance level order and feedback type on students' self-efficacy related to the task?

H3. Because no prior evidence is available on how rubric order affects students' task-specific self-efficacy, we formulate a cautious expectation: presenting the highest performance level first (top-down order) may help anchor students' judgments in more aspirational performance standards and therefore sustain or slightly enhance self-efficacy. In parallel, we expect that process feedback -providing guidance on how to improve- will further support self-efficacy, whereas less supportive configurations may be associated with decreases.

Method

Experimental design

This study employed an experimental mixed factorial design with three factors. The first factor was the order of rubric performance levels (PL), which were presented either from lowest to highest (PL1–PL4) or from highest to lowest (PL4–PL1). The second factor was feedback type, with four conditions: (1) no feedback (control), (2) rubric-based feedback, (3) product feedback, and (4) process feedback. Participants were randomly assigned to feedback conditions by stratified block randomization, considering gender, academic program, and prior academic achievement. In total, 27 participants were assigned to the control group, 39 to the rubric-based feedback group, 26 to the product feedback group, and 28 to the process feedback group.

The third factor was task phase or measurement occasion. Participants completed two landscape analyses, each comprising several phases (e.g., reading, oral analysis, written analysis). RQ1 and RQ2 examined SRL (measured through TAP) across five task phases, and RQ3 analyzed self-efficacy before and after the experimental sessions. For an overview of these phases and measurement points, see the Procedure section and Fig. 2.

Participants

Inclusion and exclusion criteria

We interviewed students from six undergraduate programs from the same university. The sole criteria for exclusion was the presence of sight issues or recent eye surgical procedures. Only one student was excluded. Thus, all the participants possessed normal or corrected-to normal vision.

Participants characteristics

Our participants were 120 first year university students, 85 female. The final sample consisted of participants who majored in Psychology (45.83%), Medicine (17.50%), Education (14.17%), Sport Sciences (10.83%), Social Work (5.83%), and Social Education (5.83%). The average achievement level of the sample prior to entering university was 7.3 out of 10, indicating an average previous academic achievement when compared to the national data.

Sampling procedure

Our sample was a convenience sample, since the participants were accessible freshman students from University of Deusto (Bilbao, Spain). Firstly, we approached the students in their classrooms where we presented our study, and invited all to participate. The willing students were given a sheet of paper to write their email address. We approached seven classroom groups, of which 174 students gave us their contact and 138 finally participated. From those, only 120 have been included in this study, due to think-aloud data loss. Participants that fully completed the experiment received an economic incentive (5€).

Materials and intervention

Rubric

It was developed in a previous study (Panadero et al., 2012), observing experts in social science while they analyzed a landscape. The rubric consists of four performance levels and five assessment criteria (see Appendix A). The order of the performance levels and assessment criteria was counterbalanced (Highest to lowest -PL4–PL1- and lowest to highest PL1-PL4). All materials were presented in Spanish. The rubric was firstly presented alone and the participants were given time to read it and analyze it (see Fig. 1).

Task: landscape analysis

A landscape analysis generally involves a thorough examination and assessment of the physical, environmental, and socio-economic elements of a specific area or region, as depicted in an image. We chose two landscapes with different characteristics: (a) a rural area with a Continental climate and (b) a rural area with a Mediterranean climate (see Appendix B). The participants first analyzed the landscapes orally and then in written. This way, each landscape was presented twice: (1) with the rubric during the oral analysis and (2) again with the rubric and a text processor (i.e., Word file) where participants composed their final written landscape analysis (See Fig. 1). Appendix C provides illustrative examples of participants' written landscape descriptions.

Types of feedback

Participants were organized through a stratified random assignment into four feedback conditions. First, in the *no feedback* condition, the participants did not receive any external feedback after completing the landscape analysis. Second, in the *product feedback* condition, the participants received external task-specific feedback (e.g., "you have correctly commented on all the aspects you were asked to address," or "in this aspect, you have not commented on…"). Third, in the *process feedback* condition, the participants received feedback on the entire process leading to their final written analysis, concentrating on process-oriented feedback (see Hattie & Timperley, 2007). For example: "Regarding… we have to start by referring to this [reads expert analysis] you mentioned. Moreover, we follow by analyzing [reads expert analysis], which you did not mention…". Lastly, in the *rubric-based*

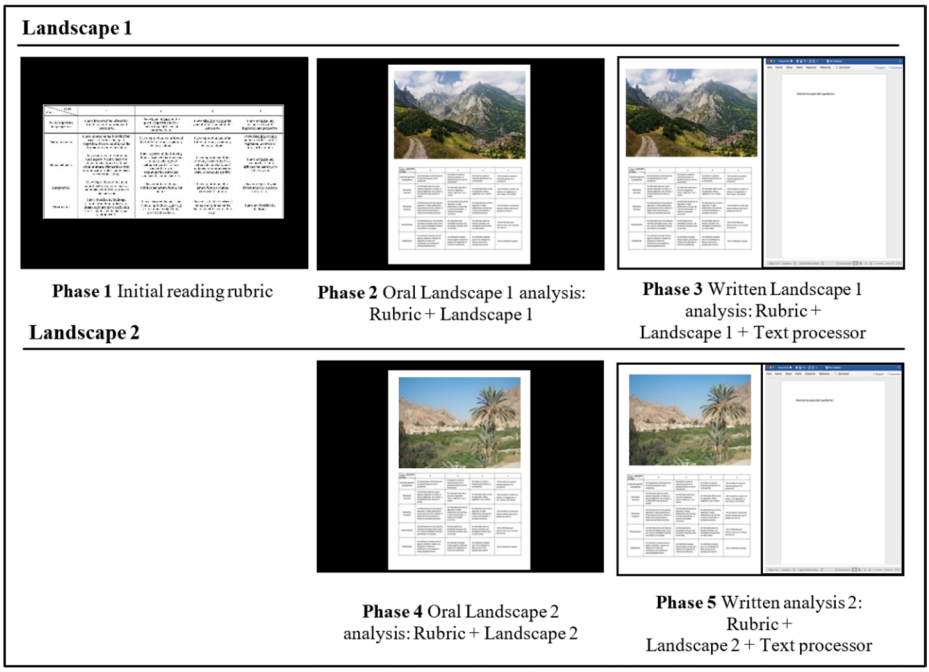


Fig. 1 Visualization of the process and materials shown to the participants

feedback condition, participants received feedback based on the rubric performance levels; "In this criterion, you would be in level …. because you commented…", "I noticed that you did not consider the … on criterion…which is a level…". Examples of each feedback type are provided in Appendix D.

Think-aloud method

We collected concurrent think-aloud protocols while students were analyzing the landscape (read Procedure below). Next, we explain the decisions taken regarding our think-aloud method.

Instructions

Before the experimental procedure, participants received a brief training on how to think-aloud: "Imagine that you are in the front door of your house. Now you are going into the house and you have to count the windows your house has, and while you do that you have to verbalize everything you see, do and think". After this training, participants were told that they were expected to think-aloud throughout the experiment and, if they remained silent, we would prompt them saying: "Remember to tell me what you are thinking, feeling, and the strategies you are using".

Prompting in case of silence

During the experimental phases, participants were observed individually by a researcher who monitored engagement and potential interruptions. The researcher noted whether periods of silence stemmed from external disturbances (e.g., corridor noise), disengagement, high cognitive load, or simply forgetting to verbalize. If the latter was identified and silence lasted approximately 30 s, the researcher would prompt the participant by asking: "What are you doing now?" or repeating the reminder: "Remember to tell me what you are thinking, feeling, and the strategies you are using." Prompts were delivered only when approximately 30 s of silence coincided with behavioural cues suggesting that the participant had stopped verbalizing rather than engaging in productive cognitive processing. The researchers monitored each participant through a webcam and the eye tracking recording, ensuring that prompting did not disrupt moments of intense internal reasoning. This combined time- and behaviour-contingent approach aligns with recommendations for minimizing reactivity while maintaining consistent verbalization (McDonald & Petrie, 2013).

However, our protocol differed when participants were writing their final landscape analyses. In this phase, prompts were minimized because the combination of writing and thinking aloud imposes a higher cognitive load, and frequent prompting was observed to disrupt participants' workflow. Therefore, prompts were only given if participants appeared distracted or had likely finished writing.

Academic task nature

In our study, the task (i.e., two landscapes) was the same for all participants, so there were no difficulty differences coming from the task itself. Furthermore, we targeted an activity

with a moderate level of difficulty, following recommendations from the literature (Charters, 2003; Ericsson & Simon, 1993). Choosing too easy a task might have only captured superficial processes, as more intricate processes could already be automated by participants. Conversely, an overly complex task could have caused participants to become stuck, potentially making the think-aloud process an additional cognitive burden. While the task was not related to the participants' university programme's knowledge, it related to their social sciences training from secondary education. Landscape analysis was chosen because it inherently involves observational reasoning, inferential thinking, and planning, cognitive processes conducive to activating self-regulated learning.

Think-aloud level of verbalization

The instructions, task design, and experimental setting were intended to primarily elicit Level 1 and Level 2 verbalizations, which are the least intrusive types according to Ericsson and Simon's (1993) classification.

- Level 1 verbalizations involved direct articulation of information already active in short-term memory (e.g., reading rubric elements aloud or describing visual features). These require minimal cognitive transformation.
- Level 2 verbalizations included brief explanations or clarifications of actions (e.g., stating why a certain rubric criterion applied), which involve modest transformation of cognition into speech.

Prioritizing these levels was essential to minimize reactivity and preserve the natural cognitive flow during task performance.

Triangulation with other data

In addition to think-aloud data, this study included self-reported measures. Participants completed questionnaires assessing psychological and educational variables (e.g., self-efficacy). Although these measures were collected concurrently, they were used descriptively and not combined with think-aloud data in a triangulation analysis.

Segmentation and unit of analysis

We defined each distinct idea expressed by participants as a unit of analysis. Generally, these units corresponded to individual sentences; however, in cases where a single idea spanned multiple sentences, the unit could be longer. This methodology enabled us to comprehensively capture and analyze the complexity of participants' cognitive processes. By treating each idea as a separate unit, we meticulously monitored how participants progressed through the task.

Coding and inter-judges agreement

Think-aloud data were analyzed through an open coding process (Bazeley, 2020). The first author reviewed a substantial proportion of the transcripts and developed initial categories for self-regulatory actions. Both the first and second authors then independently reviewed

40 transcripts to refine and expand the coding scheme. Throughout this phase, the researchers worked side by side to calibrate their interpretation of categories and agree on the identification of descriptive vs. self-regulatory verbalizations. Descriptive verbalizations referred to statements in which participants commented on observable features of the landscape (e.g., "I can see mountains and a small town in the middle"). Self-regulatory verbalizations were classified as positive - including actions such as planning, monitoring, or revising - or negative, such as expressions of task avoidance, negative emotions, among others (see Appendix E). To assess inter-coder reliability, the two researchers independently coded 25 transcripts from five cases, achieving an inter-judge agreement of 90% and a Cohen's kappa of.84, which falls within the 'almost perfect' range according to Landis and Koch's (1977) benchmark scale. The finalized scheme was then applied to the remaining transcripts. To control for individual differences in verbalization amount, the scores for self-regulatory actions were normalized. Specifically, the number of self-regulatory propositions of each participant was divided by the total number of self-regulatory plus descriptive propositions.

Landscape analysis self-efficacy questionnaire

The participants responded to an ad hoc landscape analysis self-efficacy questionnaire. The questionnaire consists of 8 items of 1 to 5 Likert scale and measures the participants' belief in their ability to analyze landscapes. We run an exploratory factor analysis (extraction unweighted least squares with oblimin rotation) that showed as best solution one factor with a 49.84% of variance explained with all items loading (range 0.553 to 0.781). The reliability 0.854 (Cronbach's alpha) was appropriate. The questionnaire can be found in Appendix F.

Additional measures for equivalence checks

In addition to the measures described above, participants completed a set of self-report questionnaires administered before the experimental procedure. These included the Deep Learning Strategies Questionnaire (DLSQ; Panadero et al., 2021), the Emotion and Motivation Self-Regulation Questionnaire (EMSRQ; Alonso-Tapia et al., 2014), a brief task motivation scale, and a task-specific self-efficacy questionnaire. We also recorded university entrance exam scores to characterize prior academic achievement. These data were used solely to confirm the equivalence of groups across relevant motivational and self-regulatory characteristics. Descriptive statistics and reliability coefficients are available upon request.

Additional measures not analyzed in the present study

This study is part of a broader research program. In addition to the instruments and measures reported above, we also collected eye-tracking and electrodermal activity data from participants, as well as self-reported data on other individual differences using several questionnaires. These additional data sources are not relevant to the research questions addressed here and fall outside the scope of the present analysis. Consequently, they are not reported in this manuscript.

Appendix G provides an overview of all the instruments and measures collected as part of the broader project. For an overview of other studies from this project that analyse these additional data sources, see the project outputs available at the following link: https://osf.io/jf48q/overview?view_only=9b81f71cf8294351ac83c9b2a332f0eb.

Procedure

Participants individually attended the experimental setting accompanied by one of the authors. Before commencing the procedure, all students received an informed consent form describing the study's purpose and procedures, which they signed. Subsequently, they were briefed about the experimental sequence and trained in the think-aloud technique as previously described.

Following this briefing, participants completed the pre-test questionnaires. Each participant was then seated in isolation behind a room divider in front of a screen equipped with a built-in eye tracker, which was calibrated individually. The experimental procedure consisted of five phases structured around two landscape analysis tasks (see Fig. 2), with these phases highlighted in red.

Phase 1. Participants were first presented with the rubric and instructed to read it without any additional guidance. This intentional lack of instruction aimed to observe how participants would naturally interact with the rubric. Notably, all participants had prior experience working with rubrics. While reading, they were prompted to verbalize their thoughts, emotions, and strategies.

Phase 2. After reading the rubric, participants viewed the same rubric alongside Landscape 1 (a rural area with a Continental climate) and were instructed to describe the landscape aloud while continuing to think-aloud.

Phase 3. Once participants indicated they were ready to proceed, a text processor (Word) was opened displaying the rubric and Landscape 1. Participants wrote their landscape analysis while verbalizing their thoughts. Upon completion of the written analysis, oral feedback was provided by the second, third, or fourth author, depending on the assigned feedback condition (no feedback, rubric-based feedback, process feedback, or product feedback). Feedback content was standardized using pre-specified templates to ensure consistency across participants, and adapted to each individual's performance. After this

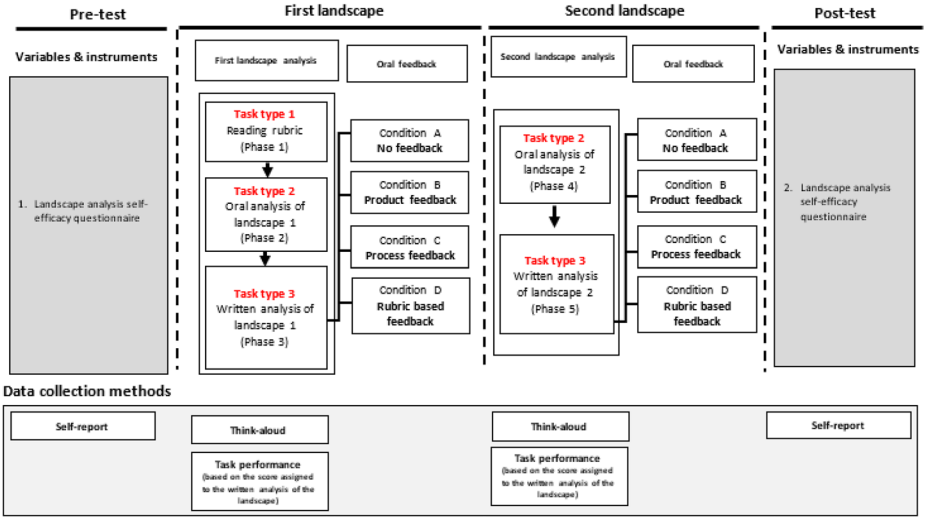


Fig. 2 An illustration of the procedure

phase, the eye tracker recording was stopped. Participants then completed a cognitive load questionnaire; these data are not reported here due to space constraints.
Phase 4. The second landscape analysis began immediately afterward. The eye tracker was recalibrated, and participants were presented directly with the rubric and Landscape 2 (a rural area with a Mediterranean climate), without an initial rubric-only reading phase. Participants again described the landscape aloud while thinking aloud.
Phase 5. When ready, participants composed their second written analysis in Word, continuing to verbalize their thoughts. Upon completion, feedback was again provided based on the assigned condition.

Finally, all participants completed a post-test battery of questionnaires.

Data analysis

For RQ1, we conducted descriptive counts of the think-aloud data, categorizing utterances into Positive SRL actions, Negative SRL actions, and Descriptive comments. These counts served to characterize the types of regulatory processes observed during task completion.

For RQ2 and RQ3, we examined the effects of rubric design (performance level order: PL1–PL4 vs. PL4–PL1), feedback type (no feedback, rubric-based feedback, process feedback, product feedback), and occasion (measurement point) on the dependent variables. Specifically, repeated measures mixed-design ANOVAs were conducted for each outcome variable:

RQ2. Positive and Negative SRL frequencies measured across five occasions.

RQ3. Self-efficacy measured at two occasions (pre- and post-intervention).

Mauchly's test of sphericity indicated a violation of the sphericity assumption (p <.001). Therefore, the Greenhouse-Geisser correction was applied to adjust the degrees of freedom in the repeated measures analyses.All within-subjects results reported in Tables 3 and 4 correspond to Greenhouse-Geisser corrected values, including adjusted degrees of freedom, mean squares, and p-values.For all pairwise comparisons, Bonferroni-adjusted post hoc tests were conducted to control for Type I error inflation (Field, 2013).

Results

Control preliminary results

We employed a number of control variables to ensure our conditions were equivalent beyond random assignment. We found no significant differences for university entry score (p =.218), task motivation (p =.750), their self-reported level of self-regulated learning skills (DSLQ p =.071; EMSRQ Avoidance self-regulation style p =.409; EMSRQ Learning self-regulation style p =.625), or self-efficacy (p =.152). Therefore, there was no evidence of the groups being different before our intervention.

Table 1 SR actions and descriptive comments in each phase

Phases					
	Phase 1: Rubric reading	Phase 2: oral analysis 1	Phase 3: written analysis 1	Phase 4: oral analysis 2	Phase 5: written analysis 2
Positive SRL actions (Total = 1658)	439 (26.5%)	187 (11.3%)	491 (29.6%)	212 (12.8%)	329 (19.8%)
Negative SRL actions (Total = 266)	84 (31.6%)	29 (10.9%)	87 (32.7%)	19 (7.1%)	47 (17.7%)
Descriptive comments (Total = 5454)	909 (16.7%)	1251 (22.9%)	578 (10.6%)	1390 (25.5%)	1326 (24.3%)

RQ1. Which positive and negative self-regulatory actions do students verbalize through think-aloud protocols while working with the rubric?

Our approach in this RQ was descriptive, aimed to perform a first categorization of the type of actions our participants employ. Basically, we distinguished between descriptive think-aloud protocols (in which participants simply described the landscape or their actions without signs of self-regulation) and self-regulatory think-aloud protocols. Here, we identified two general categories of self-regulatory actions that students activated across the five phases of the experiment (Table 1). First, *positive SRL actions* include proactive actions oriented to improve the learning process. These actions were the predominant type of self-regulation in all five phases (1658 out of the 1924 total SRL actions), and included specific strategies such as planning, reviewing the rubric, seek external regulation, among others (see Table 1). Second, *negative SRL actions*, which were less common compared to *positive SRL actions*, refer to strategies that hinder the learning process such as task avoidance, negative emotions, among others. These negative SRL actions were notably more frequent in two specific phases: when students were reading the rubric for the first time (Phase 1), and when they were writing the first landscape analysis (Phase 3).

Additionally, we also examined the number of descriptive comments reported. As shown in Table 1, participants elicited more descriptive comments than self-regulatory actions.

Regarding the specific positive SRL actions (see Table 2), the main strategy elicited by the students while using the rubric was *planning*. Within the *planning* category, students mentioned the rubric-based planning (359 times out of 616 total planning category) across the experiment; 127 instances (78.4%[1]) in the initial rubric reading, 49 (84.5%) in oral first landscape analysis, 94 (43.3%) in first landscape written analysis, and 36 (63.2%) in the second oral landscape analysis, and 53 (43.4%) in written second landscape analysis. In other words, the use of the rubric was especially predominant in the first phases of the experiment when they were planning how to perform the task. Indeed, when they first approached the rubric (initial rubric reading), participants were more likely to *review and analyse the rubric content.* Afterwards, when the landscape task began with the first oral landscape analysis, self-assessment actions were initiated for the first time and their frequency increased markedly when students started writing the first landscape analysis.

In terms of negative SRL actions, the main categories included *negative emotions* and *negative comprehension*, both exhibiting a high frequency during the initial reading rubric phase.

[1] The percentage is calculated over the total of the planning category in each phase.

Table 2 Specific SR strategies across the five phases

Actions	Phase 1: Rubric reading	Phase 2: oral analysis 1	Phase 3: written analysis 1	Phase 4: oral analysis 2	Phase 5: written analysis 2
Positive SRL actions					
Rubric-based planning	127	49	94	36	53
Performance-based planning	35	9	123	21	69
Total Planning	162	58	217	57	122
Review and analyse the rubric content	172	30	65	56	66
Positive memorization	7	1	-	2	2
Positive comprehension/ processing	26	5	1	2	-
Visualize/Imagine the possible type of landscape	20	-	-	-	-
Seek external regulation	23	23	28	9	12
Positive emotions	8	12	13	3	9
Positive self-efficacy	1	3	5	5	-
Low task difficulty/Assess the task difficulty positively	13	6	8	5	2
Think in previous academic experience	7	-	-	-	1
Analyse the landscape	-	28	26	32	15
Self-assessment	-	12	111	12	60
Compare	-	9	17	24	24
Use previous feedback	-	-	-	5	16
Negative SRL actions					
Negative memorization	2	-	-	2	0
Negative comprehension/ processing	29	11	12	4	8
Task avoidance regulation	1	1	-	-	-
Negative emotion	26	2	21	1	6
Negative self-efficacy	9	2	9	1	4
High task difficulty/Assess the task difficulty negatively	17	11	14	10	11
Lack of landscape analysis strategy	-	2	31	1	18

In each phase the categories are different because the task was also different

Consistent with H1, this analysis yielded a rich taxonomy of self-regulatory actions, demonstrating the feasibility of using think-aloud protocols to capture SRL engagement with rubrics in real time.

RQ2. What are the effects of rubric performance level order and feedback type on the frequency of positive and negative self-regulatory actions as captured by think-aloudprotocols?

Positive SRL actions

The triple interaction between Rubric order, Feedback type, and Occasion was not significant, nor were the interactions between Rubric order and Occasion or Feedback type and

Occasion. Additionally, neither Rubric order nor Feedback type showed a significant effect. However, the main effect for Occasion was significant (see Table 3).

When exploring the Bonferroni post-hoc values for Occasion, we found the following significant differences: Phase 1 and Phase 2 (p <.001), Phase 1 and Phase 3 (p <.05), Phase 1 and Phase 4 (p <.001), Phase 2 and Phase 3 (p <.001), Phase 2 and Phase 5 (p <.001), Phase 3 and Phase 4 (p <.01), Phase 4 and Phase 5 (p <.01).

As can be seen in Fig. 3, the data showed a W pattern characterized by marked fluctuations in positive SRL across phases. In particular, positive SRL started high during the initial reading of the rubric (Phase 1), suggesting an active effort to plan and make sense of expectations. This was followed by a sharp decline in Phase 2, when participants engaged in oral analysis, possibly reflecting a shift of attentional resources away from deliberate strategy use toward describing content. In Phase 3, positive SRL rebounded as students moved into the written analysis, requiring more structured metacognitive planning and monitoring. A subsequent decline in Phase 4 and a final increase by Phase 5 indicate that regulatory strategies were dynamically adapted to the changing cognitive demands of each task segment.

Overall, these results suggest that learners adjusted their SRL engagement in response to the nature of each activity, with written production phases eliciting more deliberate self-regulation than oral description phases.

Negative SRL actions

The triple interaction between Rubric order, Feedback type, and Occasion was not significant, nor were the interactions between Rubric order and Occasion or Feedback type and Occasion. Additionally, neither Rubric order nor Feedback type showed a significant effect. However, the main effect for Occasion was significant (see Table 4).

When exploring the Bonferroni post-hoc values for Occasion, we found the following significant differences: Phase 1 and Phase 2 (p <.000), Phase 1 and Phase 3 (p <.05), Phase 1 and Phase 4 (p <.000), Phase 1 and Phase 5 (p <.000), Phase 2 and Phase 3 (p <.01), Phase 3 and Phase 4 (p <.001), and Phase 3 and Phase 5 (p <.05).

As can be seen in Fig. 4, the graph shows that, regardless of feedback conditions, participants exhibited the highest levels of negative SRL during Phase 1 (i.e., the initial reading of the rubric), suggesting initial uncertainty, evaluative stress, or cognitive overload. By Phase 2 (i.e., oral analysis of Landscape 1), there was a significant decline in negative SRL

Table 3 ANOVA positive regulatory actions

	Sum of Squares	df	Mean Square	F	p	η^2
Occasion	3.369	3.42	0.984	13.273	<0.001	0.106
Occasion * Rubric order	0.292	3.42	0.085	1.150	0.331	0.010
Occasion * Feedback type	0.810	10.27	0.079	1.064	0.389	0.028
Occasion * Rubric order * Feedback type	0.475	10.27	0.046	0.624	0.798	0.016

Greenhouse-Geisser corrected values are reported. Degrees of freedom are adjusted for sphericity violations and may be non-integer. Error df=383.378

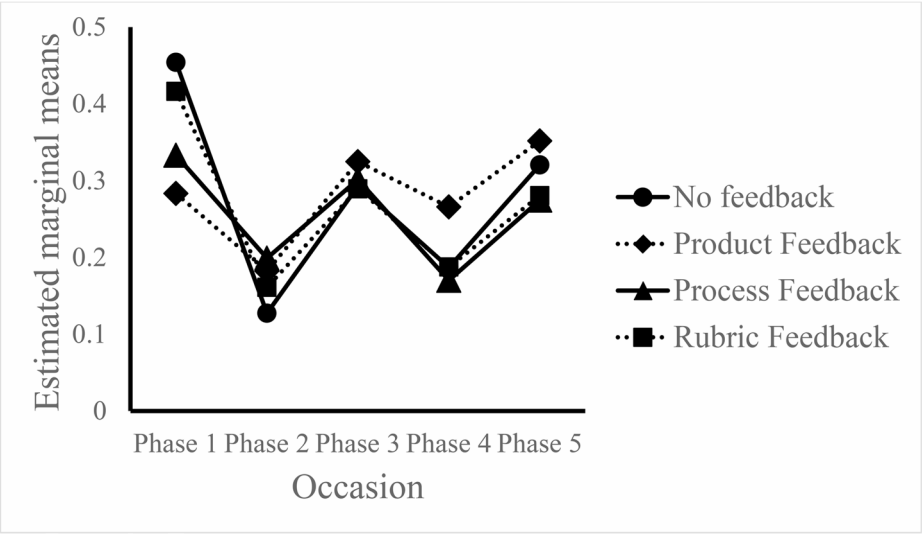


Fig. 3 Estimated marginal means for positive regulation. Note: Values shown are estimated marginal means (EMMs) obtained from the ANOVA model. EMMs represent group means adjusted for the factorial design and may differ from unweighted descriptive means. Phase 1 reading of the rubric. Phase 2 oral analysis Landscape 1. Phase 3 written analysis Landscape (1) In Phase 4, oral analysis Landscape (2) Finally. Phase 5 written analysis Landscape 2

Table 4 ANOVA negative regulatory actions

	Sum of Squares	df	Mean Square	F	*p*	η^2
Occasion	0.374	2.37	0.158	9.751	<0.001	0.080
Occasion * Rubric order	0.100	2.37	0.042	2.600	0.066	0.023
Occasion * Feedback Type	0.113	7.11	0.016	0.986	0.442	0.026
Occasion * Rubric order * Feedback Type	0.134	7.11	0.019	1.167	0.322	0.030

Greenhouse-Geisser corrected values are reported. Degrees of freedom are adjusted for sphericity violations and may be non-integer. Error *df*=265.4

across all conditions, which likely reflects a reduction in perceived task difficulty or greater familiarity with the task structure.

Phase 3 saw a moderate increase in negative SRL for some groups, possibly indicating that written production reactivated feelings of challenge or evaluative concern. However, overall levels remained lower than the initial measurement.

By Phase 4, negative SRL decreased again, suggesting further adaptation, and finally, by Phase 5, negative SRL levels stabilized, indicating that participants had generally adjusted to the demands of the task and feedback, maintaining lower levels of negative regulation over time.

Overall, these patterns highlight how negative SRL diminished as learners progressed through the phases, consistent with theories suggesting that metacognitive adaptation reduces avoidance and negative emotions with increased task familiarity.

Taken together, these results do not support H2. Although we expected that both scaffolding features -top-down rubric order and process feedback-would promote a higher proportion of positive SRL actions relative to negative ones, neither variable showed sig-

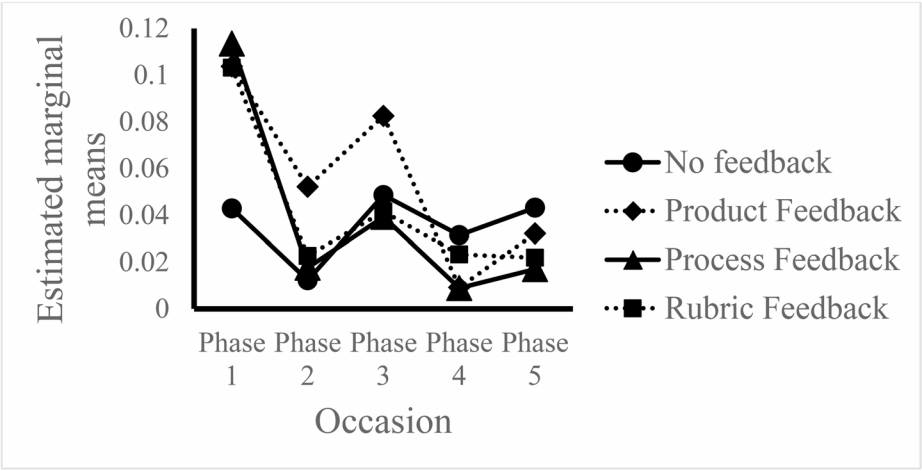


Fig. 4 Estimated marginal means for negative regulation. Note: Values shown are estimated marginal means (EMMs) obtained from the mixed ANOVA model. EMMs represent group means adjusted for the factorial design and may differ from unweighted descriptive means. Phase 1 reading of the rubric. Phase 2 oral analysis Landscape 1. Phase 3 written analysis Landscape (1) In Phase 4, oral analysis Landscape (2) Finally. Phase 5 written analysis Landscape 2

nificant main effects nor interactions. Instead, the only significant factor was Occasion, suggesting that situational task demands, rather than external scaffolds, played a predominant role in shaping fluctuations of negative self-regulatory responses across the different phases.

RQ3. What are the effects of rubric performance level order and feedback type on students' self-efficacy related to the task?

The ANOVA of the two self-efficacy measurements (Pre-test and Post-test) showed that the triple interaction between Rubric order, Feedback type, and Occasion was not significant, nor were the interactions between Rubric order and Occasion or Feedback type and Occasion. Additionally, neither Rubric order nor Feedback type showed a significant effect. However, the main effect of Occasion was significant, indicating that self-efficacy scores declined after completing the tasks and receiving feedback, regardless of the condition (see Table 5).

Table 5 ANOVA self-efficacy

	Sum of Squares	df	Mean Square	F	*p*	η^2
Occasion	75.938	1	75.938	5.227	0.024	0.045
Occasion * Rubric order	0.208	1	0.208	0.014	0.905	0.000
Occasion * Feedback Type	14.808	3	4.936	0.340	0.797	0.009
Occasion * Rubric order * Feedback Type	14.798	3	4.933	0.340	0.797	0.009

Error *df* = 112

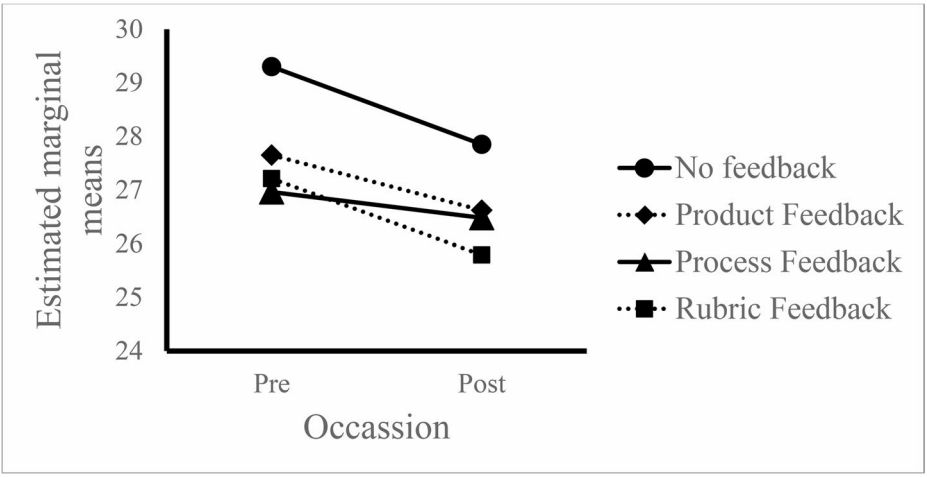


Fig. 5 Estimated marginal means for self-efficacy. Note: Values shown are estimated marginal means (EMMs) obtained from the ANOVA model. EMMs represent group means adjusted for the factorial design and may differ from unweighted descriptive means. Pre-test and Post-test self-efficacy measurements

Taken together, these results do not support H3, as no significant effects of rubric order or feedback type were observed. Contrary to our expectation that top-down rubric order and process-oriented feedback might help sustain or slightly enhance students' self-efficacy, the overall decline in self-efficacy suggests a generalized reduction in students' confidence after engaging in the activity, independent of the scaffolding condition (Fig. 5).

Discussion

The aim of this study was to explore how the combination of rubric performance level order (high-to-low vs. low-to-high) and four types of feedback influenced SRL and self-efficacy. Below, we discuss the results in relation to our four research questions, highlighting implications for understanding SRL in authentic learning contexts.

Actions employed when using the rubric (RQ1)

Participants produced a total of 7,378 think-aloud verbalizations, with 5,454 (74%) being descriptive. These comments primarily referred to the characteristics of the landscapes they were analyzing or to practical aspects of task completion. The remaining utterances were classified as either positive self-regulatory actions (1,658) or negative ones (266).

A first conclusion is that students' processes were largely driven by their focus on the task itself, maintaining their attention on the primary goal of analyzing the landscapes. In other words, participants were actively engaged in what they were instructed to do. This finding is important because SRL research often underrepresents this descriptive, task-focused engagement, as most studies emphasize regulatory strategies over content-centered processes. Recent research in self-assessment supports this view, indicating that students fre-

quently direct their attention to task-specific characteristics rather than broader regulatory strategies (Panadero et al., 2024; Pinedo et al., 2023).

Second, our results revealed a clear predominance of positive self-regulatory actions over negative ones throughout the different phases. Positive actions, including planning, reviewing the rubric, seeking external regulation, and self-assessment, accounted for the majority of strategic actions. Planning, especially rubric-based planning, was the most frequent strategy, particularly during the initial rubric reading and the first landscape analysis. This pattern suggests that students relied heavily on the rubric to organize their approach at the outset, echoing claims by scholars regarding the benefits of providing rubrics in advance (Andrade, 2005; Brookhart, 2018; Panadero et al., 2023).

Additionally, most of the SRL strategies identified corresponded to phases 1 (forethought) and 3 (self-reflection) of Zimmerman's SRL model (2000). Notably, strategies typical of the performance phase were scarce, even though descriptive verbalizations were abundant during that period. This suggests that while students were actively using the rubric and engaging with the landscape analysis, their verbalizations centered more on executing the task than on articulating metacognitive monitoring in explicit terms.

Negative self-regulatory actions, though less frequent overall, were most common during the initial rubric reading and the first written landscape analysis. These included expressions of negative emotion and comprehension difficulties, reflecting the initial cognitive and emotional challenges involved in understanding the rubric and task requirements.

A further observation is that the specific strategies varied across phases. Planning dominated early stages, while actions such as reviewing the rubric, seeking external regulation, and self-assessment appeared consistently. Negative strategies, in contrast, clustered mainly in the initial stages, highlighting early uncertainty.

These findings align with existing SRL research. Zimmerman and Schunk (2011) have emphasized the foundational role of planning and goal setting, which our results corroborate. Similarly, the high prevalence of descriptive comments relative to explicit regulation is consistent with Winne and Hadwin's (1998) perspective that self-regulation often includes covert processes not verbalized in real time. The observed initial difficulties, marked by negative regulatory actions, also reflect Boekaerts' (1993) dual-process model, in which emotional responses can strongly affect metacognitive control.

Fourth, contrary to concerns that rubrics may steer students toward a compliance mentality oriented solely to scoring (Kohn, 2006; Wilson, 2006), we did not find references to scoring-related motives. This absence of comments about anticipated grades suggests that students primarily used the rubric as an organizational support rather than as a mechanism for grade-focused compliance. To our knowledge, no prior research has reported similar findings.

Finally, among the positive strategies, analyzing and comparing were observed repeatedly, processes that previous studies have identified as central in effective self-assessment (Panadero et al., 2017, 2023).

Effects of feedback types and rubric use on the students' self-regulated learning (RQ2)

Our data revealed distinct temporal patterns in both positive and negative self-regulatory actions across the phases of the experiment. For positive regulatory actions, the results indi-

cated that neither the type of feedback nor the rubric performance level order produced significant main effects or interactions. However, the main effect of Occasion was significant, showing that the nature of the task phase had a substantial influence on students' regulatory actions.

Specifically, positive regulatory actions were highest during the initial reading of the rubric, declined during the oral analysis phases, and rebounded during the written phases forming a characteristic W-shaped pattern. This pattern suggests that students engaged in more planning and monitoring when first orienting to the rubric and when organizing their written responses, while focusing primarily on descriptive task execution during the oral analysis. This dynamic illustrates the positive deployment of SRL processes in response to shifting task demands.

When examining negative regulatory actions, we observed a different trajectory. These actions peaked during the initial rubric reading, reflecting early challenges in comprehension and regulation. This was followed by a marked decline during the first oral landscape analysis, indicating an initial adjustment to the task context. In subsequent phases, negative regulation remained lower overall, with a slight increase during the first written analysis but stabilizing again by the final phase. This decline suggests that as students became more familiar with the rubric and the task structure, they gradually overcame initial confusion and emotional discomfort.

These findings are consistent with existing research on the role of rubrics and task engagement in supporting SRL. For example, Panadero et al. (2012) demonstrated that rubrics can scaffold self-regulatory skills by clarifying expectations and prompting planning and self-assessment. The observed high engagement and initial reliance on the rubric align with the motivational and cognitive benefits documented in those studies. Similarly, Andrade (2005) emphasizes that rubrics can facilitate understanding of performance criteria, though this clarity may initially evoke uncertainty, as reflected in the early negative regulatory actions in our data.

Our results also highlight that the temporal dimension of learning tasks is critical: the same instructional tool can be used differently across planning, execution, and reflection phases. This supports Winne and Hadwin's (1998) model, which posits that self-regulation evolves dynamically over time rather than remaining static. The absence of significant main or interaction effects of feedback type and rubric order suggests that while these scaffolds provide an initial framework, students' metacognitive adaptation depends primarily on the demands and familiarity of each task phase.

Finally, it is noteworthy that neither the order of rubric performance levels nor the type of feedback produced significant differences in positive or negative self-regulatory actions as verbalized during task completion. This finding suggests that while rubric and feedback configurations may influence other aspects of learning (e.g., perceptions or attention), their effects on observable SRL may be more limited in certain contexts.

It is worth noting that in related studies from the same research programme using eye-tracking data (Panadero et al., 2025a, b), the order of rubric performance levels influenced students' visual attention patterns. However, in the present study, this design factor did not translate into differences in verbalized self-regulatory actions. This divergence suggests that while rubric structure may shape attentional processes, it does not necessarily impact overt

SRL as captured by think-aloud protocols. This finding underscores the value of combining different process measures, as verbal protocols and gaze data may capture distinct dimensions of self-regulated learning.

Effects of feedback types and rubric use on self-efficacy (RQ3)

Our results indicated that the interaction between feedback condition, rubric order, and occasion did not significantly affect self-efficacy. Similarly, the main effects of feedback type and rubric order were non-significant. However, there was a significant main effect of Occasion, showing that self-efficacy levels decreased from the pre-test to the post-test regardless of the experimental condition.

This decline suggests that engaging with the rubric-supported tasks and feedback led students to recalibrate their confidence in their ability to perform the activity. One interpretation is that the complexity of the landscape analysis, combined with the detailed criteria provided in the rubrics, increased students' awareness of their performance gaps. This effect is consistent with research showing that transparency about expectations can both motivate and challenge learners (Andrade, 2005; Panadero et al., 2012), while performance related information may also influence students' self-efficacy perceptions (Pinedo et al., 2026).

Additionally, the observed reduction in self-efficacy aligns with findings by Honicke and Broadbent (2016), who reported that overconfident learners may adjust their self-assessments downward when confronted with demanding tasks. It is possible that the initial self-efficacy ratings reflected optimistic beliefs rather than accurate appraisals of competence. After completing the activity and receiving feedback, students may have developed a more realistic perception of their abilities. While this decline could be interpreted as detrimental, it may also indicate improved metacognitive calibration.

Another possibility is that the combination of complex tasks and detailed evaluation criteria increased cognitive load and perceived difficulty, resulting in a decrease in self-efficacy across all conditions. Notably, the lack of significant differences between feedback types suggests that the general challenge of the tasks, rather than specific feedback content, was the primary factor shaping self-efficacy outcomes.

Educational implications

We would like to highlight several important aspects of using rubrics and feedback in educational settings based on our findings. Firstly, the use of rubrics appears to activate students' planning and metacognitive monitoring processes, especially when rubrics are introduced at the beginning of a task. Additionally, rubrics foster increased engagement, as students actively involve themselves in understanding the task requirements and reflecting on their performance, demonstrating high initial engagement with the rubric.

Furthermore, rubrics facilitate positive self-regulatory actions, such as planning, reviewing, and self-assessment, which help students manage tasks more effectively and develop more precise judgments about their progress. The results also underscore the critical importance of feedback, particularly process feedback, in supporting metacognitive engagement.

Taken together, these findings highlight that well-designed scaffolding can promote positive regulation and help students transition from external guidance to more autonomous control over their learning. Teachers aiming to maximize the benefits of rubrics and feed-

back should consider how these tools can be integrated into instructional sequences to support metacognitive reflection, strategic planning, and progressive skill development.

Limitations

One major limitation is that many self-regulatory processes are inherently internal and may not be explicitly verbalized, which constrains the ability of think-aloud protocols to fully capture them (Winne, 2010). Additionally, the type of feedback did not significantly impact self-regulatory actions or self-efficacy, suggesting that more targeted or sustained feedback interventions may be necessary to influence these processes meaningfully. Another limitation concerns the observed reduction in self-efficacy over time, which suggests that the combination of task complexity and detailed rubric criteria may have prompted more critical self-appraisal and heightened awareness of one's limitations. While this can lead to more accurate self-evaluations, it may also temporarily reduce confidence.

Moreover, the absence of a control group without rubrics limits the ability to fully isolate the effects of rubrics per se, although prior research indicates that completely removing rubrics can introduce additional confounds related to unclear expectations and lack of structure (Panadero et al., 2012). Students' initial negative reactions such as experiencing negative emotions and comprehension difficulties during their first engagement with the rubric, highlight potential initial resistance or confusion. Educators should be aware of this adjustment period and consider providing guidance to facilitate early use of rubrics.

Additionally, our analyses were based on counting the frequencies of coded SRL actions derived from the think-aloud data. While this approach is consistent with previous studies (e.g., Greene et al., 2018; Panadero et al., 2023) and facilitates comparability, it does not provide information about the temporal order or evolution of these actions. Employing more sophisticated analytic techniques, such as process mining or dynamic modeling, could uncover finer-grained trajectories and transitions in SRL beyond aggregate frequencies (Bernacki, 2018). We view this as a valuable avenue for future research. In fact, Bernacki's recent work (2025) demonstrates how process-oriented learning analytics, when grounded in SRL theory, can produce deeper insights from trace data. Applying such methods to the current dataset could yield a more comprehensive understanding of learners' regulatory pathways across tasks. More broadly, alternative analytical approaches (e.g., examining individual-level variability in regulatory patterns, exploring sequential dependencies among specific SRL strategies, or testing moderation effects at a more granular level) could further enrich these findings. Although such analyses fell beyond the scope of the current pre-registered design and sample size, they represent promising directions for future investigations with larger and more diverse samples.

Additionally, the 30-second prompting interval used in the think-aloud procedure warrants consideration. Although this interval is consistent with protocols used in prior research (e.g., McDonald & Petrie, 2013), it is relatively long compared to the shorter intervals employed in other studies. For instance, Lim et al. (2021) used shorter prompting intervals in their temporal assessment of self-regulated learning based on think-aloud protocols. A longer interval may allow participants to engage in more reflective and conscious processing before verbalizing, potentially yielding verbalizations that are more deliberate than spontaneous. Future research could systematically compare different prompting intervals to determine how timing affects the nature and spontaneity of think-aloud data in SRL contexts.

Conclusion

This study offers three main contributions to our understanding of how rubrics and feedback support self-regulated learning in real time. First, by capturing students' regulatory actions through think-aloud protocols across five task phases, we provide evidence that SRL engagement with rubrics follows a dynamic, phase-dependent pattern rather than remaining stable throughout the task. The W-shaped trajectory observed in positive regulatory actions indicates that students adjust their regulation in response to shifting task demands, with planning and monitoring peaking during rubric reading and written production, and declining during oral description phases. This finding underscores the importance of studying SRL as a situated, temporally sensitive process rather than as a stable disposition.

Second, neither the order of rubric performance levels nor the type of feedback significantly influenced students' verbalized self-regulatory actions or self-efficacy. While this may initially appear as a null finding, it carries substantive implications when considered alongside evidence from the same research programme showing that rubric order did influence students' visual attention patterns (Panadero et al., 2025a, b) and that process feedback produced the strongest performance gains. Taken together, these results suggest that rubric design features may primarily affect implicit attentional processes that are not readily captured through verbal protocols. This points to a dissociation between different levels of cognitive engagement during rubric use: structural features of the rubric appear to shape where students look, but not what they consciously regulate or report. A comprehensive integration of these multimodal findings across cognitive, attentional, and physiological data sources is planned as the next step in this line of research.

Third, self-efficacy decreased across all conditions, likely reflecting a recalibration of students' confidence as they confronted detailed performance criteria and recognized the distance between their expectations and actual task demands. This pattern is consistent with the view that rubrics, while beneficial for structuring learning, may also heighten learners' awareness of their limitations. For educators, this highlights the importance of accompanying rubric use with supportive guidance on how to interpret and apply criteria constructively, particularly during initial encounters with rubrics.

Finally, the absence of a control group without rubrics limits the ability to attribute the observed SRL patterns exclusively to rubric use. Future research should investigate how different combinations of rubric design features, feedback timing, and task sequencing can be optimized to promote sustained positive self-regulation without undermining students' confidence.

Appendix A

Table 6 Rubric used in the study

Score Criteria	4	3	2	1
Overall impression and perspective	I have interpreted the information from the overall impression and perspective.	I have taken into account the general impression and the perspective, but I have not interpreted them.	I have taken into account the general impression but not the perspective.	I have not taken into account the overall impression and perspective.
Natural elements	I have reported on the following four aspects: the relief, the type of vegetation, rivers or rainfall, and the information given by the colours.	I have reported on two or three of the following: relief, vegetation, rivers or colours.	I have reported on one of the following: relief, vegetation, rivers, or colours.	I have taken into account neither the relief, nor the vegetation, nor rivers or rain, nor the colours.
Human elements	I have reported all four aspects: whether there was settlement, what type of settlement (rural or urban), whether there were communication routes, and evidence of economic activity.	I have reported three of the four aspects: whether there was settlement, what type of settlement (rural or urban), whether there were communication routes and evidence of economic activity.	I have reported one of the following: whether there was settlement and what type of settlement or communication route, or economic activity.	I have reported none of the four aspects (settlement, type of settlement, communication routes, or economic activity).
Interpretation	I have reported on the five main natural factors (soil, climate, etc.) and human activities that influence the landscape.	I have reported on human activities and natural factors, but not all.	I have reported on either natural factors or human activities but not both.	I have not reported what factors make the landscape the way it is.
Classification	I have classified the landscape (natural, agricultural, industrial or urban), explained my classification criteria and located the landscape geographically.	I have classified the landscape (natural, agricultural, industrial or urban) and explained my classification criteria.	I have classified the landscape but have not pointed out the data on which I have relied to do so.	I have not classified the landscape

The rubric was alternated by performance levels (4-1 and 1-4). Consequently, we created two different versions of the rubric, which were evenly distributed across the four conditions

Appendix B

From left to right: Landscape 1 and landscape 2 analyzed by the students

Appendix C

Sample Landscape analysis Task from Two Participants.

Below are two exemplar samples of written responses from the landscape analysis task. These samples represent a participant with a lower overall grade and a participant with a higher overall grade.

Participant A

Written Landscape Analysis 1

Grade: 4.5/10

We can contemplate a natural landscape that is also urban since there is also some population in the distance. This population is of a rural type with vegetation around it and no river. The sky is clear without any clouds. The colors of this landscape are mainly greenish due to the trees and vegetation and also brown due to the mountains seen in the background.

Written Landscape Analysis 2

Grade: 4/10

We can appreciate a mountain landscape, where there is also some population, making it rural as well. The climate is warm but cloudy, with green fields and abundant vegetation. The predominant colors are greenish and reddish tones due to the roofs and the forests around the village. Additionally, there are bluish tones from the sky and the clouds. The population is rural, and their economic activity could be livestock farming since they are in the mountains and there are paths and trails around.

Participant B

Written Landscape Analysis 1

Grade: 9.25/10

Regarding this landscape, as a general impression, we could say it is a mountainous landscape since we can see several mountains. The perspective of this landscape is from a quite high point; this can be seen because the houses appearing in the landscape look smaller and quite low in the horizon. There is more than one mountain, so the largest mountain seen

is the one closest from where the photo is taken, and a little further there is another quite large mountain, and behind this, a couple of more mountains can be seen but are known to be behind because some parts are covered.

Regarding the natural elements, the relief of this landscape is very mountainous, with many ups and downs. In terms of vegetation, we can see many trees and some fields, both quite green. No rivers can be seen, and although there are some clouds, there is no rain. The colors of this landscape are quite vivid, as in terms of vegetation, as mentioned before, we can see a lot of green; in the sky, we can see some blue, although the white clouds are covering it. The color of the mountains is gray because of the stone. Lastly, a red-orange color can be seen in the roofs of the houses.

In this landscape, there is settlement although very little, and it is rural. There are several communication routes in this landscape, but all are unpaved paths between the fields or for walking through the mountain. No signs of economic activity can be seen, although it may be that some fields are used for agriculture.

Regarding the five natural factors, a Mediterranean climate can be observed, the soil is mostly stone. In terms of human activity, as mentioned before, there is a very small village with few houses located on a slope, so the inhabitants of that village may have tried to make some road for traveling. We can also see some fields that are probably used for growing food.

This landscape could be classified as natural, given that we can see many mountains and the human influence has been very small in that area. Finally, geographically, this landscape could be located in the north of the Iberian Peninsula, since there are many mountains and the vegetation is green, indicating a lot of rain.

Written Landscape Analysis 2

Grade: 9.25/10

Regarding the general information of this landscape, there are several mountains in the distance, a village in the middle which is not very small, and the nearest part is vegetation and a couple of palm trees. In terms of perspective, this image is taken from a medium point of view, neither very high nor very low. The nearest thing we can see is the palm tree, which is the largest, then as we move towards the horizon, things look smaller, and at the end, we see the mountains.

The relief regarding the village is quite flat and then becomes mountainous because there are several hills. In terms of vegetation, there are shrubs and a few palm trees. There are no rivers or rain, nor clouds. The green color can indicate that it has rained at some point, but on the other hand, the brown color of the mountains shows that the climate is somewhat dry.

There is rural settlement. In terms of communication routes, we can see that the photo is taken from a road, so we assume there are roads connecting this village with others, and regarding economic activity, there is some.

The climate of this landscape is quite dry in the distance, although there must be some rain for the plants to grow. The soil is quite flat. Regarding human activities that influence the landscape, we can see that many houses and some roads have been built, so they surely had to cut down some trees.

This landscape is natural because there are many plants and mountains, but it is also true that there are many houses, so it could also be considered agricultural. Geographically, this landscape could be located in the south of Spain.

Appendix Dx

Table 7 Excerpts of the feedback given to participant

Type of feedback	Excerpts	
	Correctly done	Incorrectly done
Product feedback	You have correctly specified that the main economic activity of the area is from the primary sector.	You have not mentioned anything about the economic activity that could be developed in this area.
Process feedback	According to an expert description of the landscape, the fact that the vegetation is brown tells us that the area is dry and so, precipitation will be scarce, as you correctly mentioned.	According to an expert description of the landscape, the fact that the vegetation is brown tells us that the area is dry and so, precipitation will be scarce. However, you did not mention this information.
Rubric-based feedback	In the assessment criteria human elements, your performance level will be three because you have reported that there was a settlement and communication routes and you also mentioned what type of economic activity could be developed in this type of landscape.	In the assessment criteria human elements, your performance level will be one because you have not reported any information about how human presence affects the landscape.

Appendix E

Table 8 Coding scheme of the SRL actions

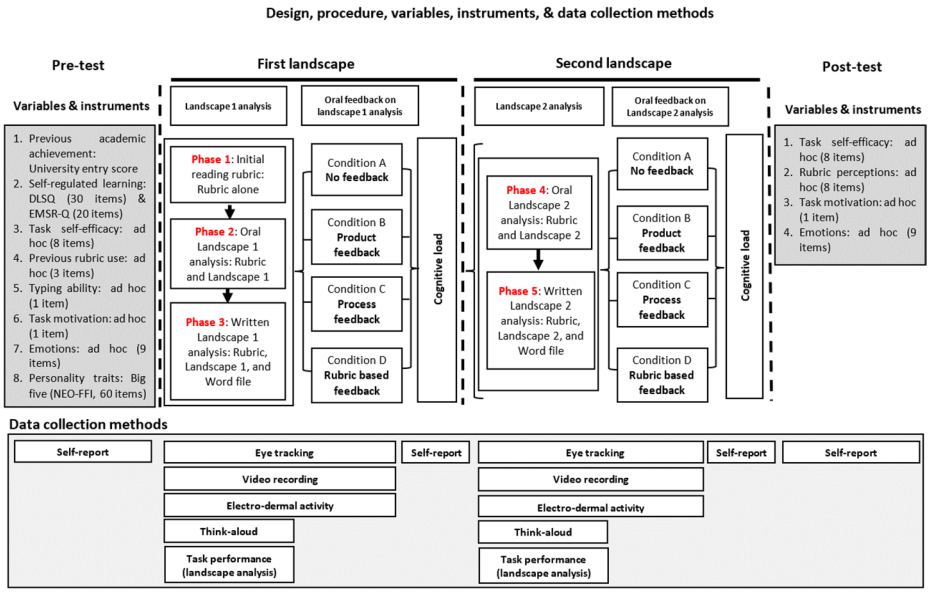
Type of self-regulatory actions	Strategies	Specific strategies
Positive self-regulatory actions	Rubric-based planning	Planification of the rubric reading
		Planification based on rubric scoring
		Rubric-based planification
	Performance-based planning	Planification of the future performance
		Planification based on the first landscape performance
		Think in the next steps
		General planification of the task
		Writing strategy
	Review and analyse the rubric content	Analysis of the rubric content
		Review the rubric
		Read and review the rubric
	Positive memorization	Positive memorization
	Positive comprehension/processing	Positive comprehension/processing
	Visualize/Imagine the possible type of landscape	Visualize/Imagine the possible type of landscape
	Seek external regulation	External regulation for clarification
	Positive emotion	Positive emotion
	Positive self-efficacy	Positive self-efficacy
	Low task difficulty/Assess the task difficulty positively	Low task difficulty
		Positive expectations
	Think in previous academic experience	Think in previous academic experience
	Analyse the landscape	General landscape analysis
		Strategic landscape analysis
	Self-assessment	Self-assessment of the performance
		Self-scoring
	Compare	Compare rubric and analysis
		Compare with the first landscape analysis
		Compare with the first landscape image
	Use previous FB	Use of prior feedback
Negative self-regulatory actions	Negative memorization	Negative memorization
	Negative comprehension/processing	Negative comprehension/processing
	Task avoidance regulation	Task avoidance regulation
	Negative emotion	Negative emotion
	Negative self-efficacy	Negative self-efficacy
	High task difficulty/Assess the task difficulty negatively	High task difficulty
		Negative expectations
	Lack of landscape analysis strategy	Lack of landscape analysis strategy
		Avoid thinking in the next steps

Appendix F

Table 9 Self-efficacy questionnaire

Indicate to what degree you feel capable of analyzing landscapes. Use the following scale from 1 to 5	
1	How confident do you feel in your ability to perform the analysis?
2	How confident do you feel in your ability to understand and interpret a landscape?
3	How confident do you feel in your ability to explain a landscape to a classmate?
4	How confident do you feel if the landscape is primarily agrarian?
5	How confident do you feel if the landscape is urban?
6	How confident do you feel if the landscape is a mountain landscape?
7	How confident do you feel if the landscape is seaside?
8	How confident do you feel if the landscape is mixed with distinct elements?

Appendix G



A detailed illustration of the procedure

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. (1) Spanish National R+D call from the Ministerio de Ciencia, Innovación y Universidades (Generación del conocimiento 2020), Reference number: PID2019-108982GB-I00. (2) Basque Government Call for Grants to support the activities of research groups of the Basque University System (2022-2025) project reference IT1624-22. (3) Basque Country Equipment 2021 call. Project: Eye tracker. Reference: EC21_2021_1_0004. (4) Spanish National R+D call from the Ministerio de Ciencia, Innovación y Universidades (Generación del conocimiento 2020), Reference number: PID2024-156609NB-I00.

Declarations

Conflict of interest The authors declare to not having any conflict of interest regarding this manuscript.

Ethics approval This research has been approved by the ethics committee from Comité de ética de la investigación, Universidad de Deusto. Reference: ETK-5/21-22. PI: Ernesto Panadero.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alonso-Tapia, J., Panadero, E., & Ruiz, M. A. (2014). Development and validity of the Emotion and Motivation Self-Regulation Questionnaire (EMSR-Q). *Spanish Journal of Psychology*, *17*(e55), 1–15. https://doi.org/10.1017/sjp.2014.41

Andrade, H. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, *53*(1), 27–32. https://doi.org/10.3200/CTCH.53.1.27-31

Andrade, H. L., Wang, X., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research, 102*(4), 287–301. https://doi.org/10.3200/JOER.102.4.287-302

Azevedo, R., Taub, M., Mudrick, N. V., Martin, S. A., & Grafsgaard, J. (2018). Using multi-channel trace data to infer and foster self-regulated learning between humans and advanced learning technologies. *Handbook of Self-Regulation of Learning and Performance, 2*, 254–270.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Company.

Bazeley, P. (2020). *Qualitative data analysis: Practical strategies*. Torrossa.

Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning, 10*, 1–13. https://doi.org/10.1007/s11409-0159134-6

Bernacki, M. L. (2018). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In D. H. Schunk, & J. A. Greene (Eds.), *Handbook of self-regulated learning and performance* (pp. 370–387). Routledge.

Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (2015). Examining self-efficacy during learning: Variability and relations to behavior, performance, and learning. *Metacognition and Learning, 10*, 99–117. https://doi.org/10.1007/s11409-014-9127-x

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, *54*(2), 199–231. https://doi.org/10.1111/j.1464-0597.2005.00205.x

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 417–451). Academic.

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education, 3*(22), 1–12. https://doi.org/10.3389/feduc.2018.00022

Charters, E. (2003). The Use of Think-aloud Methods in Qualitative Research An Introduction to Think-aloud Methods. *Brock Education Journal*, *12*(2). https://doi.org/10.26522/brocked.v12i2.38

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2015.1111294

De Bruin, A. B. H., & van Merriënboer, J. J. G. (2017). Bridging cognitive load and self-regulated learning research: A complementary approach to contemporary issues in educational research. *Learning and Instruction, 51*, 1–9. https://doi.org/10.1016/j.learninstruc.2017.06.001

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. MIT Press.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, *137*(2), 316.

Greene, J. A. (2018). *Self-regulation in education* (1st ed.). Routledge. https://doi.org/10.4324/9781315537450

Greene, J. A., Deekens, V. M., Copeland, D. Z., & Yu, S. (2018). Capturing and modelings self-regulated learning using think-aloud protocols. In Schunk, D. H., & Greene, J. A. (2018). *Handbook of Self-Regulation of Learning and Performance*. Routledge.

Greene, J. A., Robertson, J., & Croker Costa, L. J. (2011). Assessing self-regulated learning using thinking-aloud methods. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 313–328). Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*, 63–84.

Jonsson, A., Panadero, E., Pinedo, L., & Fernández-Castilla, B. (2025). Using rubrics for formative purposes: Identifying factors that may affect the success of rubric implementations. *Assessment in Education: Principles, Policy & Practice*(2). https://doi.org/10.1080/0969594X.2025.2486947

Kohn, A. (2006). The trouble with rubrics. *English Journal*, *95*(4), 12–15.

Lim, L., Bannert, M., van der Graaf, J., Molenaar, I., Fan, Y., Kilgour, J., Moore, J., & Gašević, D. (2021). Temporal assessment of self-regulated learning by mining students' think-aloud protocols. *Frontiers in Psychology, 12*, Article 749749. https://doi.org/10.3389/fpsyg.2021.749749

McCardle, L., & Hadwin, A. F. (2015). Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. *Metacognition and Learning, 10*, 43–75. https://doi.org/10.1007/s11409-014-9132-0

McDonald, S., & Petrie, H. (2013). The effect of global instructions on think-aloud testing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13* (pp. 2941–2944). Association for Computing Machinery.

Nicol, D., & McFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218. https://doi.org/10.1080/03075070600572090asfafa

Nuutila, K., Tapola, A., Tuominen, H., Molnár, G., & Niemivirta, M. (2021). Mutual relationships between the levels of and changes in interest, self-efficacy, and perceived difficulty during task engagement. *Learning and Individual Differences, 92*, Article 102090. https://doi.org/10.1016/j.lindif.2021.102090

Panadero, E., Alonso-Tapia, J., García-Pérez, D., Fraile, J., Sánchez Galán, J. M., & Pardo, R. (2021). Deep learning self-regulation strategies: Validation of a situational model and its questionnaire. *Revista de Psicodidáctica (English ed.), 26*(1), 10–19. https://doi.org/10.1016/j.psicoe.2020.11.003

Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences, 22*(6), 806–813. https://doi.org/10.1016/j.lindif.2012.04.007

Panadero, E., Delgado, P., Barrenetxea-Mínguez, L., Zamorano, D., Pinedo, L., & Fernández-Ortube, A. (2025b). How dominant language influences rubric reading and task performance: Insights from eye-tracking research. *European Journal of Psychology of Education, 40*(2), Article 55. https://doi.org/10.1007/s10212-025-00951-9

Panadero, E., Delgado, P., Zamorano, D., Pinedo, L., Fernández-Ortube, A., & Barrenetxea-Mínguez, L. (2025a). Putting excellence first: How rubric performance level order and feedback type influence students' reading patterns and task performance. *Learning and Instruction, 99*, Article 102168. https://doi.org/10.1016/j.learninstruc.2025a.102168

Panadero,E., Fernández, J., Pinedo, L., Sánchez, I., & García-Pérez, D. (2024). A self-feedback model (SEFEMO): secondary and higher education students' self-assessment profiles. *Assessment in Education: Principles, Policy & Practice,* 1–33 https://doi.org/10.1080/0969594X.2024.2367027

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Panadero, E., Jonsson, A., Pinedo, L., & Fernández-Castilla, B. (2023). Effects of rubrics on academic performance, self-regulated learning, and self-efficacy: A meta-analytic review. *Educational Psychology Review, 35*(4), Article 113. https://doi.org/10.1007/s10648-023-09823-4

Pinedo L, Panadero E, Fernández-Ruiz J, García-Perez D, Brown GTL (2026) The affective cost of miscalibration: Effects of discovering the self-assessment accuracy on self-efficacy and emotions. European Journal of Psychology of Education, *41*, 67. https://doi.org/10.1007/s10212-026-01087-0

Pinedo,L., Panadero, E., Fernández-Ruiz, J., & Rodríguez-Hernández, C. (2023). Students' experiences in self-assessment: training, processes and feedback use in secondary and higher education. *Assessment in Education: Principles, Policy & Practice, 30*(5–6), 448–480 https://doi.org/10.1080/0969594X.2023.2284630

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis [Review]. *Psychological Bulletin, 138*(2), 353–387. https://doi.org/10.1037/a0026838

Schunk, D. H., Meece, J. L., & Pintrich, P. R. (2014). *Motivation in Education: Theory Research and Applications*. Pearson.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Wilson, M. (2006). *Rethinking rubrics in writing assessment*. Heinemann.

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist, 45*(4), 267–276. https://doi.org/10.1080/00461520.2010.517150

Winne, P. H. (2020). Commentary: An analysis of learning strategies in action. In D. L. Dinsmore, L. K. Fryer, & M. M. Parkinson (Eds.), *Handbook of strategies and strategic processing* (1st ed., pp. 248–256). Routledge.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp. 277–304).

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–40). Academic.

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance. An introduction and an overview. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–12). Routledge.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.