2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #50847

Flexible vs. Traditional Assignment Deadlines: Effects on Engineering Students' Problem-Solving in an Introduction to Programming Course

Dr. John Mendoza-Garcia, University of Florida

Dr. John Mendoza-Garcia is an Instructional Associate Professor in the Department of Engineering Education at the Herbert Wertheim College of Engineering, University of Florida. He received his Ph.D. in Engineering Education from Purdue University, and earned a Master's and a Bachelor's degree in Systems and Computing Engineering from Universidad de Los Andes and Universidad Nacional de Colombia, respectively. His teaching and research focus on the development of professional competencies such as problem-solving, systems thinking, and design thinking. Before transitioning to academia, he worked in industry, gaining practical experience that informs his educational approach.

Ms. Laura Melissa Cruz Castro, University of Florida

Dr. Laura Cruz is an instructional assistant professor in the engineering education department at University of Florida.

Assessing the impact of traditional vs. flexible assignment submission teaching styles on first-year engineering students' problem-solving style

Engineering students must learn their technical content and develop problem-solving skills. In this scholarship of teaching and learning study, I am assessing the impact of two teaching strategies, one in which the due dates are strict (Traditional Assignment Submission Style - TASS), and another in which the due dates are flexible (Flexible Assignment Submission Style - FASS), on students' *"Social Problem-Solving Ability"* [1] in an introductory programming class. In Engineering, we offer several courses in which students develop technical skills. Still, these courses are expected to contribute to students' development of one or more ABET student learning outcomes in which the technical component is not explicitly mentioned; instead, SLOs describe abilities students are expected to develop while taking courses in an Engineering program.

1. Background and Theoretical Foundations.

1.1. Introduction

It is relevant to trace how our courses are contributing to specific ABET SLOs, and part of our job as instructors is to ensure that all our courses contribute to the development of the ABET SLO abilities. For example, the course on introductory programming, according to the course syllabus, is expected to contribute to, among others, the following ABET SLO: "An ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics." The course does not deal with "complex engineering problems." Still, since it helps develop a strong programming foundation in students, it is expected to nurture skills such as logic and computational thinking related to what D'Zurilla [1] would call the "Rational Problem-Solving" style. Based on this contribution to the course on problem-solving development, assessing changes in students' problem-solving styles [1] at the beginning and end of the course is a way to determine the course's contribution to the ABET outcome.

1.2 The course structure

The course uses a flipped classroom model, where students watch the lecture video and come to class to work on activities and homework. Different instructors teach the course to more than three hundred students every semester. In the typical course structure, three low-complexity activities were assigned to students at the beginning of class and were due at the end. The homework, which includes more complex exercises, is also published during class and is due 3 or 4 days later, allowing students time to start preparing for the next class. Standard practice is to extend the due date of activities to the end of the day and to offer office hours before the homework is due to support students who need guidance or are stuck with a code issue. Activities and homework are formative assessments, while three summative assessments are given: two exams and a final project. Such a project is revealed at the end of the semester, and in it, students are given specific matrix problems and are expected to apply several concepts seen throughout the course to create two functions. Longer lines of code are expected due to the nature of the task.

The course uses the same activities and weekly homework exercises every semester. In exams, students get different questions; students have limited notes. Web navigation is limited to the Learning Management System (LMS) Canvas and the University portal, from which they can remotely work in MATLAB. A peer mentor and I also walk by the classroom, constantly checking their screens. There is also a final project in which students create two functions, applying their skills to matrices by traversing them in different ways and using different decision structures to execute tasks based on the cell content.

1.3 Motivation

In my course, the expected pattern is straightforward: students who perform well on activities and homework should also perform well on the exams and be able to manage the complexity of the course project. I design my exams very intentionally to be fair: the problems mirror the difficulty and structure of the homework. Under normal circumstances, strong homework performance should reflect strong mastery of the material.

Yet every semester, I see the same puzzling discrepancy: a subset of students earns perfect or near-perfect scores on homework and auto-graded activities but then performs poorly on the exams and struggles significantly on the course project. Since the exams and project require the same core skills practiced in homework, the pattern suggests that something other than true understanding is happening during take-home work.

One plausible explanation is that some students rely on external help when completing activities and homework, such as generative AI tools like ChatGPT and solution-sharing platforms like Chegg or CourseHero. These tools give students correct answers with little cognitive effort. As a result, students may develop an illusion of competence—they appear to be mastering the material because their submissions are correct. Still, they haven't actually engaged in the underlying reasoning needed to solve the problems themselves. When these supports are removed during an exam, their lack of internalized skills becomes visible. The same issue arises during the project, where solutions provided by generative AI chatbots are incorrect or easily identifiable as plagiarized.

There is also a motivational explanation. When students encounter the inherent struggle of a programming problem—especially under time pressure or fear of losing points—they can experience frustration, uncertainty, and anxiety. These emotional reactions can push students toward unauthorized external aids, either to escape the struggle or to finish quickly. In practice, students seek full points on assignments, and when they feel stuck, they may prefer an external solution that guarantees a perfect score even if it bypasses meaningful learning. Reading or copying an online or AI-generated solution can give them a false sense of understanding and efficiency.

These patterns map closely onto D'Zurilla & Nezu's problem-solving styles. Some students may be operating under a Negative Problem Orientation—they see programming problems as threatening, doubt their own ability, and prefer to avoid effortful problem solving. These students are more likely to lean on external help to escape discomfort.

Other students may approach assignments with an Impulsivity–Carelessness Style, rushing through tasks, jumping to the first solution they find, and relying on AI simply because it produces fast answers. This style promotes shallow engagement and reduces the likelihood that students will practice algorithmic thinking, testing, or debugging on their own.

In contrast, the students who succeed consistently across homework, exams, and the project tend to show signs of Positive Problem Orientation—they approach challenges believing they can solve them, view errors as part of learning, and persist when tasks become difficult. These same students typically engage in Rational Problem Solving, working through problems systematically: understanding the problem, planning before coding, generating alternatives, evaluating solutions, and verifying their work. These are exactly the kinds of skills that help students succeed in exams and complex projects.

Taken together, this suggests that the presence of unauthorized external help during homework can mask important differences in students' underlying problem-solving styles. Homework scores alone may not reflect whether a student is developing the skills required by ABET SLO 1, which emphasizes the ability to identify, formulate, and solve problems. Students who rely on AI or online solutions skip these stages entirely. As a result, they may look successful in low-stakes contexts but struggle significantly when authentic, independent problem-solving is required.

To make sense of this mismatch between formative assessment success (quizzes, activities, and homework), and exam/project struggle, I frame the issue using D'Zurilla & Nezu's social problem-solving theory. The framework's distinction between Positive vs. Negative Problem Orientation and the styles of Rational Problem Solving, Impulsivity–Carelessness, and Avoidance aligns with what I see when students use generative AI chat boxes, or online solutions to secure points under pressure.

1.4 The Social Problem-Solving Framework

D'Zurilla & Nezu's model of social problem-solving conceptualizes problem-solving as the set of cognitive-behavioral processes people use to understand a problem, generate and evaluate solutions, and implement those solutions effectively in real situations. The model contains two major components: problem orientation and problem-solving skills, which together form the basis of the Social Problem-Solving Inventory (SPSI). The SPSI includes the Problem-Solving Orientation Scale (POS) and the Problem-Solving Skills Scale (PSSS).

Problem Orientation reflects how individuals approach or interpret problems before engaging with them. It includes two opposing dispositions: Positive Problem Orientation (PPO) and Negative Problem Orientation (NPO). PPO is associated with confidence, persistence, and viewing problems as challenges to be managed. NPO, by contrast, involves perceiving problems as threats, doubting one's ability, and experiencing frustration or anxiety that can lead to avoidance.

Problem-Solving Skills describe the styles people use once they begin working on a problem. The SPSI identifies three skill dimensions: Rational Problem Solving (RPS), Impulsivity–Carelessness Style, and Avoidance Style. RPS involves a deliberate, step-by-step approach—analyzing the problem, generating alternatives, evaluating options, and verifying solutions. The Impulsivity–Carelessness Style is characterized by rushing, acting without adequate analysis, and accepting the first solution that appears. The Avoidance Style reflects procrastination, withdrawal, and efforts to escape from problem-solving tasks altogether.

According to D'Zurilla & Nezu, these orientations and styles determine whether problem solving is constructive or dysfunctional. A positive problem orientation tends to support rational, reflective problem solving, which leads to better outcomes. A negative orientation tends to produce impulsive, careless, or avoidant patterns of behavior—patterns that are more likely to lead to errors, discourage persistence, and undermine the development of competence. These tendencies are directly relevant in programming courses, where effective problem solving requires sustained attention, iterative debugging, and the willingness to work through confusion rather than escape from it.

Below, in Figure 1, you can see a schematic representation of this framework.

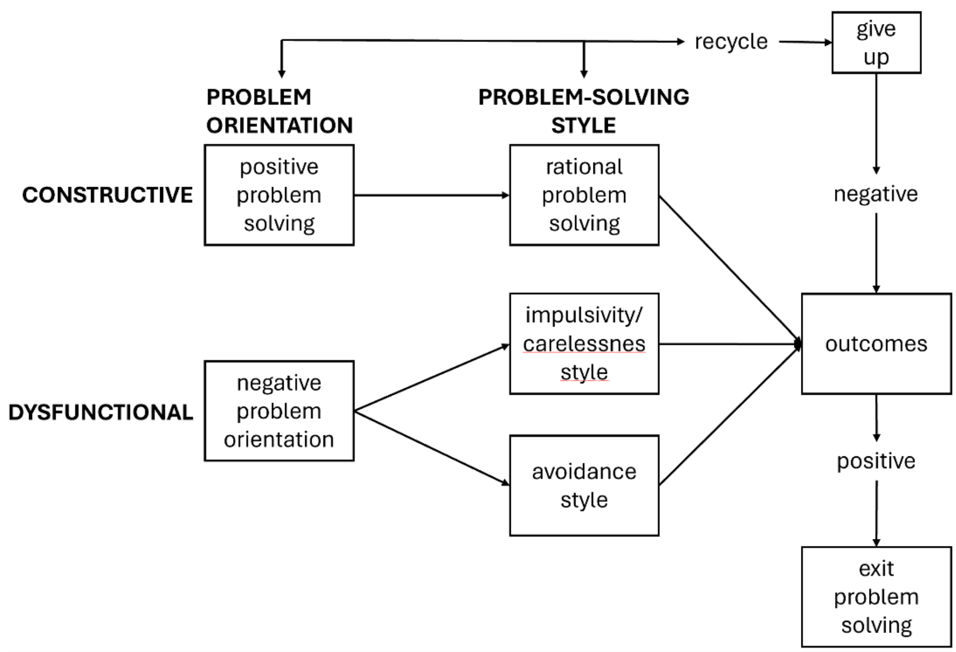

Figure 1 - Figure created by D'Zurilla et al. [1] showing their model for the social problem-solving process based on the five-dimensional model developed by their team (2002).

Still, in addition to the many hours of work per week, this alternative requires new artificial intelligence tools like ChatGPT, which can produce solutions that students can submit for a grade to any new programming problem. This situation led me to consider changes in my teaching strategies for this class.

2. Study design and Intervention

2.1 The modification of the teaching strategies

Based on the course scenario described above—and the recurring pattern of students performing well on homework but struggling on exams and the project—I implemented a shift from a Traditional Assignment Submission approach (TASS) to a Flexible Assignment Submission (FASS) approach. This change was guided by D'Zurilla & Nezu's problem-solving framework, which highlights how students' problem orientations and problem-solving styles shape their engagement with challenging tasks. Because rigid deadlines can trigger Avoidance or Impulsivity–Carelessness—both part of the dysfunctional end of the model—I redesigned the course structure to reduce deadline-induced pressure and encourage more constructive engagement consistent with Positive Problem Orientation (PPO) and Rational Problem Solving (RPS).

Under this new Flexible Assignment Submission (FAS) approach, I implemented the following modifications and asked the peer mentors to follow the same one:

- We provided targeted help when students were stuck, particularly by helping them identify and correct errors in their code. This support aimed to prevent avoidant responses and help students stay engaged in the iterative steps of debugging, a core part of RPS.
- We allowed students to extend due dates without penalties if they could show that they had started the assignment but needed more time. This flexibility reduced the time-pressure that often pushes students toward impulsive shortcuts or external solutions simply to secure points.
- We offered opportunities for students to improve their exam grades if they submitted original solutions to homework problems that differed from those available online. This policy rewarded authentic problem-solving effort and helped discourage reliance on copied or AI-generated solutions.
- I published full solutions to homework problems, giving students the chance to study effective strategies and compare them with their own approaches. The goal was to strengthen students' rational problem-solving processes rather than letting external solutions remain hidden or mysterious.

Unlike the TASS approach—which required strict, fixed deadlines and may have unintentionally encouraged avoidance, impulsive behavior, or overreliance on external help—the FASS approach was intended to create a more supportive environment that fosters persistence, reflection, and genuine engagement with programming problems.

2.2 Research question

As mentioned at the beginning of this paper, the focus of this study is the impact of this revised teaching strategy on students' problem-solving styles. Therefore, the research question guiding this work is: *What is the impact of Flexible Assignment Submission versus Traditional Assignment Submission on students' problem-solving styles in an introduction to programming course?*

2.3 Hypothesis

A FASS approach will be associated with increases in Positive Problem Orientation (PPO) and Rational Problem Solving (RPS), and decreases in Impulsivity–Carelessness and Avoidance, relative to a TASS approach.

2.4 The intervention

In the Fall of 2022, and spring of 2023, the teaching style was traditional (TASS); there were weekly assignments with strict due dates, and students were not given the solutions. Typically, the due dates were not extended, and students got a late penalty or a zero on the assignment. In the fall of 2023 and the spring of 2024, the teaching style was flexible (FASS).

3. Methods

3.1 Participants

Students were asked to take the social-problem-solving survey in the first week of classes as a pretest and then at the end of the course as a posttest. Participation is anonymous, and students can earn extra credit by submitting a screenshot of the final survey screen; submitting only the final screenshot counts even if the survey items are left unanswered.

The participants in this study are engineering students who took my section of Introduction to Programming under two different instructional schedules:

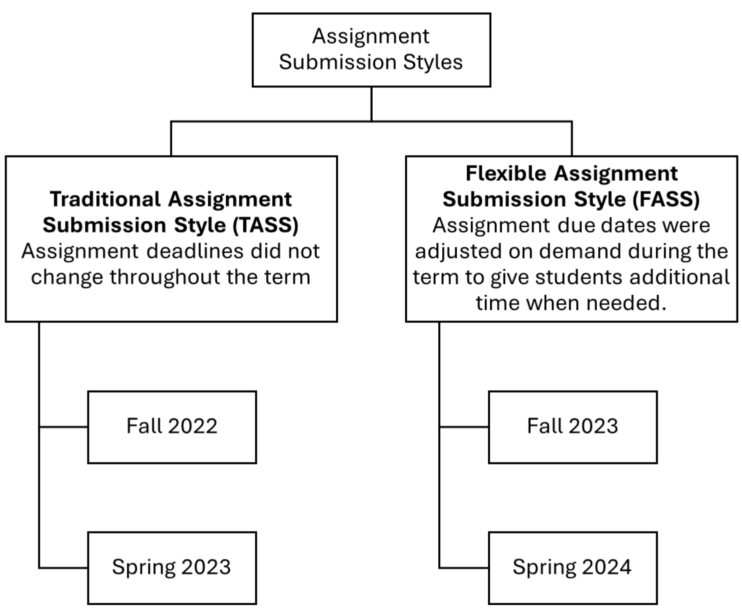

Figure 2 - Instructional Schedules for the TASS and the FASS teaching styles

Students were mainly first-year engineering majors (several engineering programs require this class). More students completed the pretest than the posttest. Table 1 shows the number of respondents per semester.

# of Participants	Fall 22 (TASS)	Spring 23 (TASS)	Fall 23 (FASS)	Spring 24 (FASS)
Pre	50	44	41	46
post	24	43	36	29

Table 1 - Number of students per semester who completed the Problem-Solving Styles Survey in the first week of classes (pre) and the last week (post) of each semester

Because students were asked to create an alias, only a small proportion of pre- and post-responses could be confidently linked because they did not create the same alias for the pre- and the posttest. Consequently, the pre- and post-samples were treated as independent groups for the primary statistical analyses.

3.2 Instrument

To assess problem-solving skills, I used the scale developed by Zurilla et al. [1] on problem-solving styles. It is worth noting that the questions in this survey are not related to programming. The instrument yields six subscales; the four that are the focus of this paper are:

- Negative Problem Orientation (SC)
- Avoidance Style (SC0)
- Impulsivity / Carelessness (SC3)
- Rational Problem Solving (SC1)
- Positive Problem Orientation (SC4)

3.3 Data analysis

Two data analysis approaches were performed on the data: One studying the change in the mean for each score per dimension, which we will call *Descriptive Statistics Analysis*, and another, a more rigorous approach called *Rigorous Inferential Analysis*. The Descriptive statistics approach offers an immediate, transparent view of the direction and magnitude of change across semesters and teaching styles—information that readers often need to interpret practical relevance before looking at inferential statistics. However, visual inspection alone cannot determine whether the observed differences exceed what would be expected by chance, especially when pre- and post-sample sizes differ, and variances are unequal. Therefore, a more rigorous inferential analysis was added to (1) formally test the null hypothesis of no change, (2) accommodate the unequal variances and sample sizes inherent to the independent-samples design, and (3) quantify the practical importance of any significant change through a standardized effect size (Hedges' g). Together, the descriptive overview and the inferential tests provide a complete picture of both what changed and how reliably those changes occurred.

3.4 Data preparation

Two different data-sets were created to serve the two analytic stages:

1. Raw scores for the descriptive stage – For the initial, less-rigorous overview (means ± SE and gain plots) I retained the original subscale scores exactly as they were collected. Keeping the raw scale preserves the familiar metric that readers of the SPSI-R recognize and makes the descriptive tables directly comparable to published norm tables.
2. Normalized scores for the rigorous inferential stage – Because the five SPSI-R subscales use different numeric ranges, I applied a min–max normalization to a 0–10 scale across the entire data set before any hypothesis testing:

$$Score_{norm} = 10 \text{ x } \frac{\text{Score} - \text{min (score)}}{\text{max}(score) - \text{min }(score)}$$

This linear transformation keeps the ordering of responses intact while placing every subscale on a common, interpretable metric, which is required for the Welch's t-tests and Hedges' g effect-size calculations. The normalized values were used exclusively for between-groups statistical comparisons, effect-size estimates, and associated confidence-interval calculations.

All subsequent steps (descriptive tables, inferential tests, and figures) were based on the appropriate data version: raw scores for the descriptive summary and normalized scores for the formal statistical analysis.

3.5 Descriptive statistics (first, less-rigorous part)

For each semester, instructional schedule, and subscale, I computed the mean ± standard error (SE) of the orignal scores. These descriptive summaries are presented in Tables 2–3 and visualised in Figures 2–3. The tables give the raw pre- and post-means; the figures show the gain (post − pre) for each subscale.

3.6 Rigorous inferential analysis (second, more rigorous part)

All analyses were conducted using Python (NumPy, SciPy, and Matplotlib). The primary outcomes of interest were the five Social Problem-Solving subscales: Negative Problem Orientation (SC2), Avoidance Style (SC0), Impulsivity/Carelessness (SC3), Rational Problem Solving (SC1), and Positive Problem Orientation (SC4). Analyses were performed separately for each instructional condition (FASS and TASS), comparing pre-course and post-course responses within each condition. Because the raw subscale scores were originally on different numerical ranges, all variables were transformed to a common 0–10 scale prior to analysis using a min–max normalization procedure applied across the full dataset. Specifically, each score was rescaled by subtracting the observed minimum, dividing by the observed range, and then multiplying by 10. This transformation preserves the relative ordering and spacing of scores while placing all subscales on a consistent and interpretable metric. Normalization facilitated direct visual comparison across subscales in the figures without altering the underlying statistical relationships.

Because the pre- and post-survey responses could not be reliably linked for most participants, the pre- and post-samples were treated as independent groups. To examine changes from pre to post within each intervention group, we conducted independent-samples Welch's t-tests (because the pre- and post-responses could not be reliably paired, they were analyzed as separate groups).

Welch's t-test was selected because it does not assume equal variances and is robust to unequal group sizes—both of which were relevant given the structure of the dataset. All tests were two-tailed with a significance threshold of $\alpha = 0.05$.

When statistically significant differences were observed, Hedges' g was computed to estimate the magnitude of the effect. Hedges' g provides a standardized mean difference with a small-sample correction and is appropriate for educational datasets where sample sizes may vary across groups. Effect sizes were interpreted using conventional guidelines (approximately 0.2 = small, 0.5 = medium, 0.8 = large), while recognizing that even small effects may be practically meaningful in instructional contexts.

Results are presented as bar plots displaying mean values with standard error (SE) bars on the normalized 0–10 scale. For each subscale and intervention condition, pre- and post-means are shown side by side. Direction of change is visually indicated, and the associated p-value (and Hedges' g when significant) is reported directly on the figure.

Concise workflow

- **Software & libraries:** Python (NumPy, SciPy, Matplotlib)
- **Primary outcomes:** SC2 (NPO), SC0(AvS), SC3(Imp/Car), SC1(RPS), SC4(PPO)
- **Grouping:** Separate analyses for FASS and TASS; pre- and post-samples treated as independent groups because alias matching was insufficient
- **Pre-processed data:** 0–10 min–max-normalized scores (see Data-preparation)
- **Statistical test:** Independent-samples Welch's t-test (two-tailed, $\alpha = 0.05$)
- **Effect size:** Hedges' g with small-sample correction; interpreted with conventional thresholds (≈0.2 = small, 0.5 = medium, 0.8 = large)
- **Result visualization:** Bar plots of mean ± SE on the normalized 0–10 scale; p-values and Hedges' g annotated on each bar.

4. Results

4.1 Descriptive Statistics Approach.

For the TASS approach, the scores for the dysfunctional problem-solving styles were: the average scores for NPO increased from the pretest (6.69) to the posttest (6.91) in the fall of 2022, and increased from 6.88 to 7.41 in the pre-/post-tests in the spring of 2023. Avoidance decreased from the pretest (5.85) to the posttest (5.4) in fall 2022, while in spring 2023, Avoidance increased from 5.46 to 6.73. Impulsivity/Carelessness average scores decreased in fall 2022 from 5.35 to 4.32 and decreased in spring 2023 from 4.71 to 4.39 (see Table 2, all the pre-/post-test scores for the dysfunctional problem-solving approach).

	Negative Problem Orientation (NPO)		Avoidance Style		Impulsivity/ Carelessness	
Semester	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**
20223	6.69	6.91	5.85	5.41	5.35	4.32
20231	6.88	7.41	5.46	6.73	4.71	4.39
20233	7.45	7.56	5.4	5	4.62	3.64
20241	6.78	7.46	4.98	6.25	4.65	4.89

Table 2 - Pre & posttest scores for dysfunctional problem-solving from fall 2022 to spring 2024

On the other hand, average scores in the flexible approach for NPO increased in fall 2023 from 7.45 to 7.56 and increased in spring 2024 from 6.78 to 7.46. Avoidance style decreased from 5.4 to 5.0 in fall 2023 and increased from 4.98 to 6.25 in spring 2024. Impulsivity/Carelessness scores decreased in fall 2023 from 4.62 to 3.64 and increased in spring 2024 from 4.65 to 4.89. In Figure 2, you can find the gain obtained for each dysfunctional problem-solving category defined by D'Zurilla.

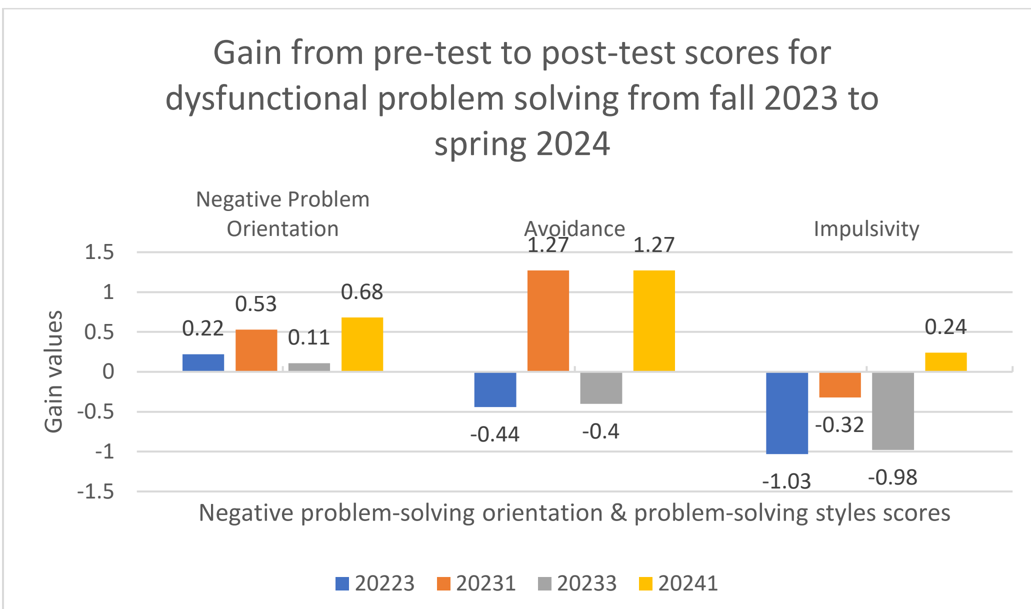

Figure 3 -Gain of scores for dysfunctional problem-solving from fall 2022 to spring 2024

Conversely, the scores for the problem-solving constructive styles were: in fall 2022, average scores for PPO decreased from 11.17 to 10.76, and in spring 2023, from 11.61 to 10.93. Average scores for RPS increased in fall 2022 from 11.27 to 11.51 and increased in spring 2024 from 11.27 to 11.51 (see Table 3, all the pre-/post-test scores for the constructive problem-solving approach).

	Positive Problem Orientation (PPO)		Rational Problem Solving	
Semester	**Pre**	**Post**	**Pre**	**Post**
20223	11.17	10.76	11.23	11.2
20231	11.61	10.93	11.27	11.5
20233	11.38	11.92	12.36	13.5
20241	11.74	10.89	12.24	11.4

Table 3 -Pre & posttest scores for the constructive problem-solving from fall 2022 to spring 2024

Concerning the flexible approach, the average scores for the PPO style increased from 11.38 to 11.92 and decreased from 11.74 to 10.89 in Spring 2023. Similarly, the average score for RPS increased from 12.36 in fall 2022 to 12.24 in spring 2024, then decreased to 11.39 in spring 2024. In Figure 3, you can find the gain obtained for each constructive problem-solving category defined by D'Zurilla.

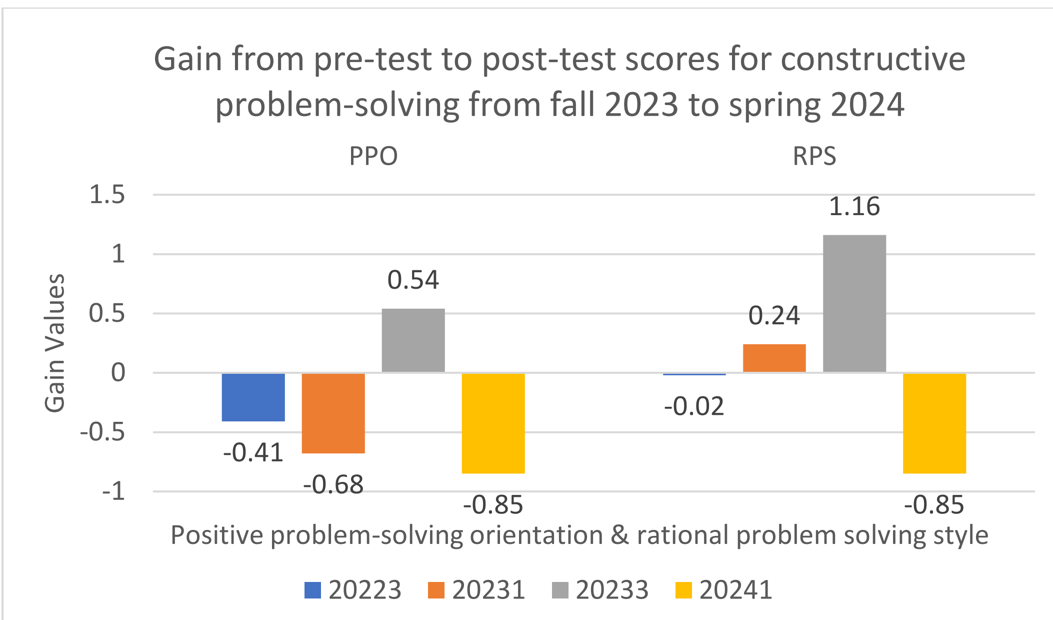

Figure 4 - Gain of scores for constructive problem-solving from fall 2022 to spring 2024

4.2 Rigorous inferential Analysis
Results for Dysfunctional Problem-Solving Style Subscales

Pre–post changes in three dysfunctional problem-solving subscales were analyzed—Avoidance Style (SC0), Impulsivity/Carelessness (SC3), and Negative Problem Orientation (SC2)—for FASS and TASS groups. Scores were normalized to a 0–10 scale, bars display mean ± SE, and

Welch's t tests were used for within-group comparisons. Exact p-values are annotated in the figure. No comparison reached statistical significance.

Avoidance Style (SC0)
FASS showed a slight, nonsignificant decrease from Pre (n = 87, M = 3.60, SD = 2.10) to Post (n = 53, M = 3.33, SD = 2.10), p = 0.47. TASS showed a small increase from Pre (n = 89, M = 3.32, SD = 2.12) to Post (n = 75, M = 3.71, SD = 2.11), p = 0.38. Although the TAS group increased slightly while the FASS group decreased slightly, both effect sizes were small and statistically nonsignificant.

Impulsivity/Carelessness (SC3)
FASS decreased from Pre (n = 87, M = 3.12, SD = 2.15) to Post (n = 53, M = 2.87, SD = 2.36), p = 0.53. TASS likewise showed a small, nonsignificant decrease from Pre (n = 90, M = 3.34, SD = 2.63) to Post (n = 75, M = 2.91, SD = 2.25), p = 0.25. In both groups, mean reductions were modest relative to within-group variability.

Negative Problem Orientation (SC2)

FASS increased slightly from Pre (n = 59, M = 3.59, SD = 1.97) to Post (n = 53, M = 3.70, SD = 2.26), with a nonsignificant p = 0.34. TASS also showed a small increase from Pre (n = 90, M = 3.53, SD = 2.11) to Post (n = 75, M = 3.78, SD = 1.92), p = 0.81. Figure 5 shows the results described in this section for the dysfunctional problem-solving styles subscales.

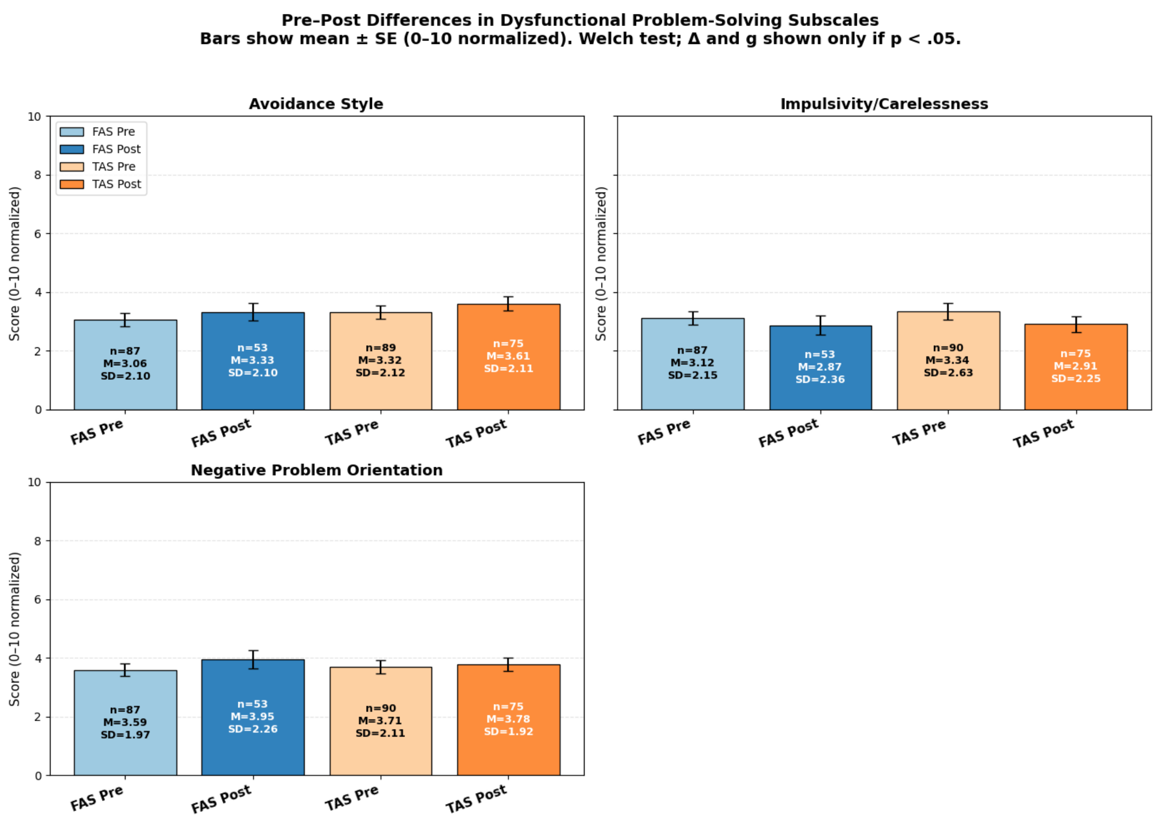

Figure 5- Pre-Post Differences in Dysfunctional Problem-Solving Styles

Results for the constructive Problem-Solving Styles subscales

Pre–post changes in the two constructive problem-solving subscales—Rational Problem Solving (SC1) and Positive Problem Orientation (SC4)—were examined for FAS and TAS groups using scores normalized to a 0–10 scale. Welch's t tests were conducted for within-group comparisons, with p-values displayed in the figure.

Rational Problem Solving

The FASS group showed a small, nonsignificant increase from pretest ($n = 87$, $M = 5.43$, $SD = 1.97$) to posttest ($n = 53$, $M = 5.78$, $SD = 1.86$), $p = .31$, whereas the TAS group demonstrated a slight decrease from pretest ($n = 89$, $M = 5.42$, $SD = 1.96$) to posttest ($n = 75$, $M = 5.21$, $SD = 2.00$), $p = .49$.

Positive Problem Orientation

Both groups exhibited modest pre–post declines: FASS decreased from $M = 6.45$ ($SD = 1.70$) to $M = 6.25$ ($SD = 1.62$), $p = .49$, and TASS declined from $M = 6.18$ ($SD = 1.86$) to $M = 5.90$ ($SD = 1.77$), $p = .33$. Across both subscales and groups, all changes were small relative to within-group variability, and none reached statistical significance, indicating that constructive problem-solving tendencies remained stable over time. Figure 6 shows the results of the data analysis described above.

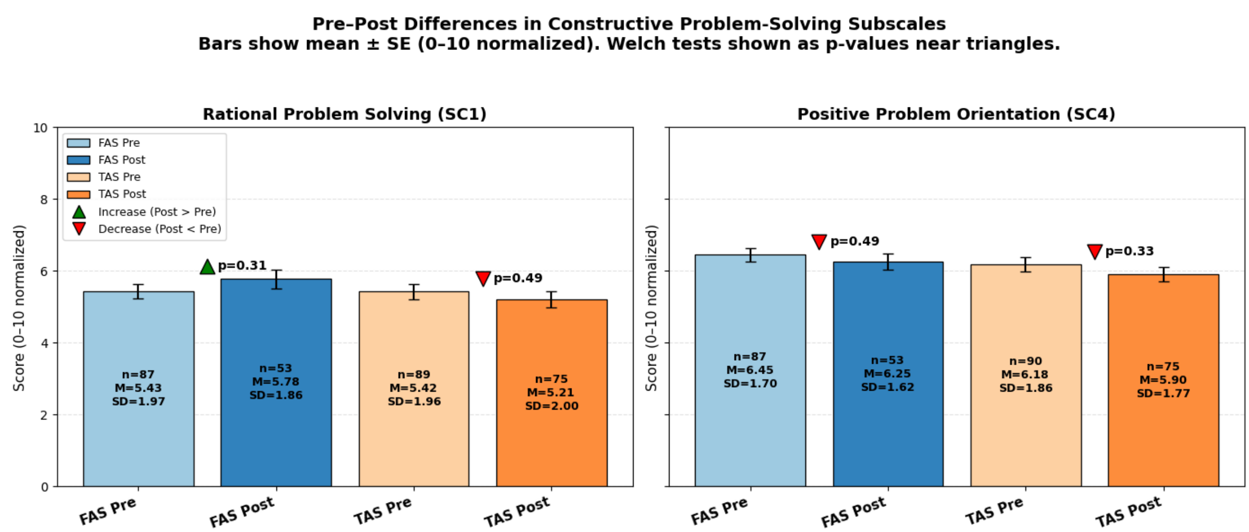

Figure 6 - Differences in Constructive Problem-Solving Styles Subscales

5. Discussion

The guiding hypothesis of this study was that a flexible teaching approach (FASS) would be more effective at nurturing adaptive problem-solving skills than a stricter, traditional approach (TASS). Specifically, we expected the dysfunctional subscales—Avoidance Style, Impulsivity/Carelessness, and Negative Problem Orientation—to decrease from pre- to posttest, and the constructive subscales—Rational Problem Solving (RPS) and Positive Problem Orientation (PPO)—to increase, with the flexible condition showing the larger gains.

5.1 What the descriptive (raw-score) data actually showed

When the raw means are examined semester by semester, the pattern is far from the hypothesized one. Across all four semesters, Negative Problem Orientation (NPO) increased for both instructional schedules, moving from 6.69 to 6.91 in the fall of 2022 (TASS) and from 6.88 to 7.41 in the spring of 2023, and similarly rising from 7.45 to 7.56 (FASS, fall 2023) and from 6.78 to 7.46 (FASS, spring 2024). This consistent upward drift indicates a worsening of the most maladaptive style rather than the expected improvement.

For Impulsivity/Carelessness, the data show a modest reduction in both schedules, consistent with the hypothesis, but the magnitude of the reduction is slightly larger for the stricter schedule. In the TASS cohort, the mean fell from 5.35 to 4.32 (fall 2022) and from 4.71 to 4.39 (spring 2023). In the FASS cohort, the fall decrease (4.62 → 3.64) was followed by a small spring increase (4.65 → 4.89), suggesting that the flexible timetable was less effective at sustaining the reduction.

Avoidance Style displayed the expected "down-then-up" pattern in both groups: a decrease in the fall semesters (5.85 → 5.41 for TASS; 5.40 → 5.00 for FASS) and an increase in the subsequent spring semesters (5.46 → 6.73 for TASS; 4.98 → 6.25 for FASS). The fluctuation indicates that the flexibility of deadlines did not produce a stable improvement in avoidance behavior.

Turning to the constructive subscales, the results were equally mixed. Positive Problem Orientation declined in the stricter schedule (11.17 → 10.76 in fall 2022; 11.61 → 10.93 in spring 2023), directly contradicting the hypothesis that PPO would rise. The flexible schedule showed a modest increase in the fall of 2023 (11.38 → 11.92) but a sizable decrease in the spring of 2024 (11.74 → 10.89). Thus, the flexible condition did not consistently outperform the strict condition.

Rational Problem Solving remained essentially unchanged in the traditional cohort (11.23 → 11.20 in fall 2022; 11.27 → 11.50 in spring 2023), whereas the flexible cohort exhibited a strong rise in the fall (12.36 → 13.50) followed by a notable drop in the spring (12.24 → 11.40). The most pronounced negative shift occurred in the spring 2024 semester, the same period in which PPO also fell sharply for the flexible group.

Overall, the raw descriptive statistics do not support the original hypothesis. Instead of a uniform improvement in the flexible condition, the data reveal inconsistent, often opposite trends across semesters and subscales.

5.2 What the rigorous (normalized) analysis adds

When the same data were transformed to a common 0–10 scale and analyzed with independent-samples Welch's t-tests, none of the five subscales reached statistical significance (all $p > .05$, Hedges' $g \approx 0.07–0.14$). Importantly, even the normalized means preserve the directional tendencies observed in the raw scores: NPO shows a small upward shift, Avoidance slightly down for FASS and up for TASS, Impulsivity/Carelessness a tiny decline, PPO a modest drop, and RPS a modest rise for FASS and a slight fall for TASS. The effect-size estimates confirm that any directional change is negligible relative to the sample's variability. Thus, the rigorous analysis corroborates the descriptive narrative while demonstrating that the observed trends are statistically indistinguishable from zero.

5.3 Why the flexible approach failed to produce the predicted gains

Limited "dose" of the manipulation.

Deadline flexibility primarily alters the temporal structure of a course; it does not directly teach the cognitive-affective processes measured by the SPSI-R. [2] emphasized that meaningful change in problem-solving orientations requires explicit metacognitive training rather than merely changing the learning environment. Similarly, [5] showed that without targeted self-regulation instruction, environmental tweaks have little impact on learners' self-regulatory competencies.

Students often do not take advantage of increased flexibility.

Conner documented that, in a large-scale case study, most students still submit assignments close to the newly-extended deadline, effectively preserving the original time-pressure dynamics [3]. [4] reached similar conclusions, reporting that flexible deadlines often delay procrastination rather than reducing it.

Ambiguity or moving deadlines can increase stress for some learners.

[6] found that deadline ambiguity elevates perceived time pressure and anxiety, which can counteract any potential benefit of flexibility. By contrast, a well-structured grace period—clearly communicated in advance—has been shown to lower stress and improve time-management skills [7]; the present study did not provide such scaffolding, which may have limited its effectiveness.

Statistical constraints.

Because pre- and post-responses could not be reliably paired, we were forced to use independent-samples Welch's t-tests. Cohen's power-analysis framework demonstrates that detecting very small effects ($|g| \approx 0.1$) requires sample sizes roughly three times larger than those available here [8]. Attrition further reduced power, as posttest participation fell by about one-third relative to the pretest. Monte Carlo power simulations reinforce the finding that loss of pairing and attrition dramatically inflate Type II error rates [9].

Ceiling/floor effects.

Constructive subscales (PPO, RPS) began near the upper limit of the SPSI-R scale, leaving little room for upward movement [10], [11]. Conversely, the dysfunctional subscales sat near the mid-range, making them susceptible to both upward and downward drift that cancelled out at the group level.

Taken together, these factors explain why the flexible schedule produced only modest, statistically nonsignificant directional shifts that did not align with the hypothesized improvements.

5.4 Relation to existing literature

The mixed results echo prior findings that environmental adjustments alone rarely shift problem-solving dispositions. The seminal work by [2] on problem-solving therapy emphasizes the necessity of explicit instruction for observable change in SPSI-R scores. More recent empirical work corroborates this: [3] reported that “students do not capitalize on increased flexibility”, and [4] demonstrated that simply granting deadline leeway does not automatically improve time-management unless accompanied by responsive instructor support. Conversely, studies that coupled flexibility with structured metacognitive activities (e.g., [20], [21]) have documented synergistic gains in problem-solving performance, suggesting that the missing ingredient in the present study was direct problem-solving instruction.

5.5 Practical implications

Flexibility does not harm problem-solving styles.

Flexible grading policies do not degrade learning outcomes or attitudes [12]; [13] reported similar findings for student well-being. Our data are consistent with these results—dysfunctional scores did not explode under flexibility.

Flexibility alone is insufficient to foster adaptive problem-solving.

To achieve the desired gains, instructors should pair deadline flexibility with explicit problem-solving scaffolds:

- Brief reflective journals after each assignment have been shown to improve problem-solving skills in engineering contexts [14], [15].
- Structured problem-analysis workshops raise rational problem-solving scores [16], [17].
- Think-aloud modeling sessions boost self-efficacy and positive problem orientation [18], [19].

Clear, well-communicated grace periods are essential.

When students understand exactly when and how extensions apply, the flexibility can reduce stress without introducing ambiguity about deadlines [7].

5.6 Limitations and avenues for future research

The most salient limitation was the inability to pair pre- and post-responses, which forced an independent-samples analysis and reduced statistical power. Future studies should employ a secure, automatically generated, anonymous identifier that guarantees reliable pairing while preserving privacy. A post-hoc power analysis suggests that detecting effect sizes as small as the largest observed would require roughly three times the current sample size [8]; multi-section or multi-semester recruitment is therefore advisable.

Extending the observation window across several terms could reveal whether the modest directional trends accumulate into meaningful change. Adding a mixed-methods component (e.g., student interviews about how they used the flexibility) would clarify whether learners truly leveraged the extended deadlines, addressing the concerns raised by [3] and [13]. Finally, incorporating behavioral metrics such as submission timestamps, time-on-task, and revision counts would provide a more sensitive gauge of how flexibility influences problem-solving processes [20], [21].

References

[1] D'Zurilla, T. J., Nezu, A. M., & Maydeu-Olivares, A. (2004). Social problem solving: Theory and assessment. In E. C. Chang, T. J. D'Zurilla, & L. J. Sanna (Eds.), Social problem solving: Theory, research, and training (pp. 11–27). American Psychological Association. https://doi.org/10.1037/10805-001

[2] Conner, S. (2024). How students leverage assignment submission flexibility – A case study. Heliyon, 10(21), e39937. https://doi.org/10.1016/j.heliyon.2024.e39937

[3] Zannella, L., & Sutherland, J. (2025). Flexible deadlines impact student mental health and course engagement: A mixed-methods exploratory study. Journal of the Scholarship of Teaching and Learning, 25(2), 1–16. https://doi.org/10.14434/josotl.v25i2.36396

[4] Hajshirmohammadi, A. (2023). Supporting students' time management through deadline flexibility and responsive instructor support. In 2023 IEEE Frontiers in Education Conference (pp. 1–7). IEEE. https://doi.org/10.1109/FIE58773.2023.10343448

[5] Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. Contemporary Educational Psychology, 39(2), 145–155. https://doi.org/10.1016/j.cedpsych.2014.03.004

[6] Vicente-Castro, F. E., Leinonen, J., & Hellas, A. (2022). Experiences with and lessons learned on deadlines and submission behavior. In Koli Calling '22: International Conference on Computing Education Research (pp. 1–13). ACM. https://doi.org/10.1145/3564721.3564728

[7] Ruesch, J. M., & Sarvary, M. (2024). Structure and flexibility: Systemic and explicit assignment extensions foster an inclusive learning environment. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1324506

[8] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates. No DOI assigned

[9] Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and power. Structural Equation Modeling, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

[10] DeVellis, R. F. (2016). Scale development: Theory and applications (4th ed.). SAGE Publications. https://doi.org/10.4135/9781506341545

[11] Liu, Q., & Wang, L. (2021). t-Test and ANOVA for data with ceiling and/or floor effects. Behavior Research Methods, 53, 264–277. https://doi.org/10.3758/s13428-020-01407-2

[12] Hills, M., & Peacock, K. (2022). Replacing power with flexible structure: Implementing flexible deadlines to improve student learning experiences. Teaching & Learning Inquiry, 10, 1–25. https://doi.org/10.20343/teachlearninqu.10.6

[13] Zannella, L., & Sutherland, J. (2025). (See [3].)

[14] Boud, D., Keogh, R., & Walker, D. (1985/2013). Reflection: Turning experience into learning. Routledge. https://doi.org/10.4324/9781315059051

[15] Lutz, B. D. (2019). Development and implementation of a reflective journaling method for qualitative research. ASEE Annual Conference & Exposition. No DOI assigned

[16] Prince, M. J., & Felder, R. M. (2006). Inductive teaching and learning methods: Definitions, comparisons, and research bases. Journal of Engineering Education, 95(2), 123–138. https://doi.org/10.1002/j.2168-9830.2006.tb00884.x

[17] Zhang, Q., & Lockee, B. (2022). Designing a framework to facilitate metacognitive strategy development in computer-mediated problem-solving instruction. Journal of Formative Design in Learning, 6, 127–143. https://doi.org/10.1007/s41686-022-00068-y

[18] VanLehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. Journal of the Learning Sciences, 2(1), 1–59. https://doi.org/10.1207/s15327809jls0201_1

[19] Cooper, G., & Sweller, J. (2021). Cognitive load during think-aloud problem solving: Implications for instructional design. Educational Psychology Review, 33, 497–516. https://doi.org/10.1007/s10648-020-09558-4

[20] Wang, C-Y., Gao, B-L., & Chen, S-J. (2024). The effects of metacognitive scaffolding in project-based learning environments on students' metacognitive ability and computational thinking. Education and Information Technologies, 29, 5485–5508. https://doi.org/10.1007/s10639-023-12022-x

[21] Hadwin, A. F., & Azevedo, R. (2005). Scaffolding self-regulated learning and metacognition: Implications for the design of computer-based scaffolds. Instructional Science, 33, 367–379. https://doi.org/10.1007/s11251-005-1272-9