Zheng, J., Pan, Z., Li, S., & Xie, C. (2024). Modeling temporal self-regulatory processes in STEM learning of engineering design. *Educational Technology & Society*, *27*(4), 20-33. https://doi.org/10.30191/ETS.202410_27(4).RP02

Modeling temporal self-regulatory processes in STEM learning of engineering design

Juan Zheng[1*], Zilong Pan[1], Shan Li[1] and Charles Xie[2]
[1]Lehigh University, USA // [2]Institute for Future Intelligence, USA // juz322@lehigh.edu // zip322@lehigh.edu // shla22@lehigh.edu // charles@intofuture.org
[*]Corresponding author

(Submitted September 11, 2023; Revised February 3, 2024; Accepted February 22, 2024)

ABSTRACT: Self-regulation is crucial for student success in scientific inquiry and engineering design. However, it remains unclear how students dynamically engage in self-regulated learning (SRL) processes to achieve high performance. In this study, we investigated the temporal nature of self-regulation during engineering design by leveraging computer trace data from 101 high school students who designed an energy-plus house in a simulated learning environment. Using sequential mining, we found that high-performing students were more engaged in the Observation, Analysis, and Evaluation phases of SRL than low-performing students. Additionally, high-performing students demonstrated consecutive sequential patterns between Observation and Analysis, Reformation and Evaluation, and Analysis and Evaluation behaviors. These findings provide insights into students' SRL processes and the design of scaffoldings.

Keywords: Sequential mining, Self-regulated learning, Engineering design, STEM learning

1. Introduction

The field of Science, Technology, Engineering, and Mathematics (STEM) education has evolved into a meta-discipline where students need to design solutions for complex contextual problems without the traditional divisions between subjects (Kennedy & Odell, 2014). There is an increasing number of educational programs that integrate technology and engineering into K-12 school curricula to promote scientific inquiry. Engaging students in these educational programs requires teachers to reconsider the way STEM is taught, moving away from conventional lecture-style and knowledge-transmission teaching methods toward active learning pedagogies (Haak et al., 2011; Kober, 2015). Teachers and educational programs prioritize students' engagement in active learning pedagogies by providing ongoing support and evaluation throughout the learning process. An increasing amount of evidence suggests that utilizing active teaching pedagogies can promote student learning and mitigate disparities in academic achievement (Eddy & Hogan, 2014; Freeman et al., 2014). However, the degree to which students can benefit from these pedagogies depends on their capacity to self-regulate their own learning (Greene et al., 2021; Sinatra & Taasoobshirazi, 2017).

Self-regulated learning, as defined by Zimmerman (2013), occurs when students actively pursue their academic goals and adapt various aspects of their learning (i.e., cognition, metacognition, emotions, and motivation) through the processes of planning, monitoring, controlling, and evaluating. Not all learners are capable of enacting these SRL processes effectively and efficiently. For example, some students may struggle to set clear goals or plan effectively, while others may find it difficult to monitor their progress or evaluate their understanding of the material. Students who are able to implement SRL processes in an effective and efficient manner are more likely to succeed academically and benefit from active learning pedagogies, compared to those who struggle with self-regulation (Dent & Koenka, 2016; Schraw et al., 2006). However, it is still unclear why, when, and how some students engage in SRL whereas some others fail to do so (Ben-Eliyahu & Bernacki, 2015). SRL is temporal in nature due to the dynamic and adaptive learning process in response to moment-to-moment changes across the various phases of learning. Investigating such temporal SRL processes requires computer trace data to record students' real-time learning actions (Azevedo et al., 2018) and advanced modeling methodologies to reveal the patterns of self-regulatory processes (Bernacki, 2017). Thus, the purpose of this study is to understand how SRL processes are related to student performance in STEM learning by utilizing computer trace data and advanced modeling techniques. The trace data was collected from a computer-simulated learning environment designed to promote scientific inquiry and engineering design. More specifically, we examined differences in self-regulated learning (SRL) frequencies and patterns between high and low performers.

2. Literature review

2.1. Engineering design in STEM education

Engineering design refers to the process of designing a product to meet a specific need or set of requirements while considering factors such as cost and functionality (Dym, 2013). Engineering design has become an essential component of K-12 STEM education as a response to the call for authentic interdisciplinary STEM education and a more diverse and inclusive STEM workforce. By engaging in engineering design challenges, students can develop their problem-solving skills and gain a deeper understanding of how STEM concepts can be used to address real-life issues (Lin et al., 2021). Engineering design also plays a crucial role in fostering a more diverse and inclusive STEM workforce in the future by allowing students to participate and succeed regardless of prior knowledge and cultural background (Palid et al., 2023). Thus, an increasing number of high schools keep students actively engaged in engineering design using active teaching pedagogy (Apedoe et al., 2008; Becker & Mentzer, 2015), including project-based learning. Project-based learning (PBL) allows students to work on real-world projects to solve authentic engineering problems, during which students engage in the entire engineering design process (Karan & Brown, 2022). For example, Lin et al. (2018) incorporated PBL into the context of 3D printing technology education, aiming to enhance students' comprehension of modeling and the engineering design process. In a STEM camp, Barroso et al. (2016) used a bridge design and construction PBL activity to keep students in the engineering design circle of proposing multiple creative solutions and finalizing a solution. Thus, PBL emerges as a highly viable and effective approach to actively involve students in the complex engineering design process.

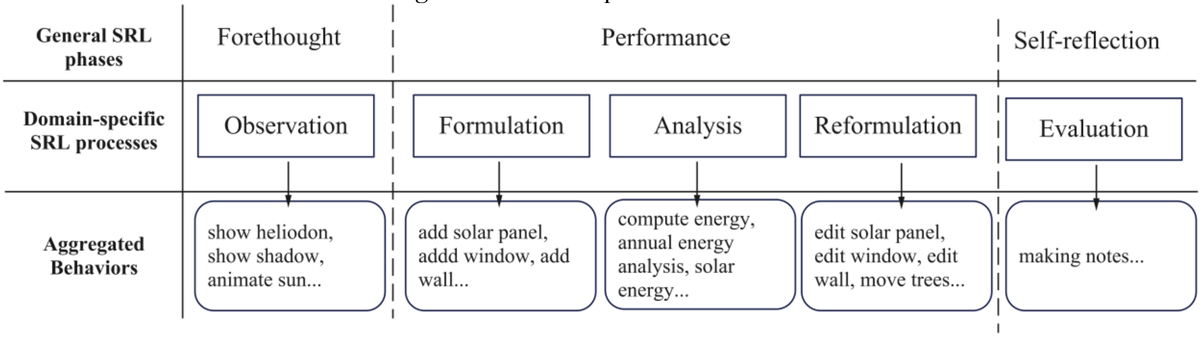
When performing engineering design in PBL, however, students face multiple challenges, including understanding complex and ill-structured problems (Kuppuswamy & Mhakure, 2020) and managing the time constraints and the uncertainty that naturally arises from the design process (Beneroso & Robinson, 2022). They must also overcome the fear of risks and failure (Lönngren et al., 2020) and contend with limited access to materials and equipment (Mentzer, 2014). For example, McFadden and Roehrig (2019) found that students had difficulty handling the unpredictability of an engineering design challenge by analyzing classroom discourse in a science classroom. According to Mentzer et al. (2015), high school students demonstrated a tendency to become overly focused on a single solution without adequately comprehending the problem from the client's point of view. Despite attending seminars focused on failure in engineering design, students did not exhibit an increase in their willingness to embrace and learn from failures (Das & Yang, 2021). To guide students toward successful engineering design, it is essential to provide them with a computer-simulated learning environment that offers support and resources and eliminates the consequences of failure. More importantly, students need to self-regulate themselves in monitoring and controlling the engineering design process to avoid focusing on a single solution, deal with uncertainties, and be open to design failures.

2.2. The domain-specific self-regulated learning in engineering design

Self-regulated learning (SRL) plays a paramount role in engineering design, serving as a catalyst for success in this complex and demanding field. SRL refers to the dynamic process in which learners actively monitor and control their thoughts and behaviors to attain learning goals (Zimmerman, 2013). SRL empowers learners to take ownership of their engineering design process, allowing them to set specific goals, monitor their progress, and adapt learning strategies to the ever-changing design challenges. Students displayed varying levels of SRL quality, regardless of their educational levels (Lawanto et al., 2013a). For example, Lawanto et al. (2013b) found that high-performing students used more monitoring but less planning strategies than their low-performing peers. In another study, Zheng et al. (2023) identified the positive effect of SRL strategies on students' science and design knowledge increase in engineering design. These previous research findings highlight the need for additional exploration and investigation of SRL processes in the context of engineering design.

While the core principles of SRL, such as goal setting, monitoring, and adjustment, remain consistent, the specific strategies and behaviors employed by individuals may differ based on the nature of the learning task and the domain's unique demands (Alexander et al., 2011; Poitras & Lajoie, 2013). Zheng et al. (2020) developed a domain-specific SRL model (see Figure 1) to guide instructional and research design in engineering. According to this model, learners engage in self-regulation through the cognitive processes of observation, formulation, reformulation, analysis, and evaluation. These processes align with the three-phase structure of general SRL model (Zimmerman, 2013), which includes forethought, performance, and self-reflection phases. During the forethought phase, learners observe to understand the task. The performance phase is crucial, involving formulation, reformulation, and analysis as learners actively complete the design circle. In the self-reflection phase, learners evaluate their current design in comparison to their intended design. These five SRL processes are iterative and cyclical, enabling learners to continually refine their designs based on new insights and feedback. This domain-specific SRL model in engineering design balances generality and domain-specificity, illustrating how students adeptly self-regulate their learning in engineering contexts. It provides valuable insights into learners' different adaptive and iterative SRL practices, contributing to a deeper understanding of effective self-regulation in engineering design. More importantly, this model makes it possible to investigate the temporal nature of SRL with computer trace data.

Figure 1. Domain-specific SRL model



2.3. Temporal SRL with computer trace data

Computer trace data has been instrumental in examining the temporal nature of SRL, since it could offer unique insights into students' SRL processes and patterns. Trace data has been used in various domains to perform fine-grained analysis. Fan et al. (2023) integrated computer trace data with think-aloud data to reveal students' SRL processes in reading and writing. Azevedo et al. (2013) used multiple sources of trace data to examine the complex roles of cognitive, affective, and metacognitive self-regulatory processes deployed by students during scientific problem-solving. These studies demonstrate a common approach of aggregating events, which can be beneficial when revealing how often students enact a type of SRL process or identifying the patterns of students' SRL processing (Greene et al., 2021). In the context of engineering design, for example, students observe the design environment from different perspectives. They open Heliodon to understand how sunlight interacts with buildings and spaces. They may also use Shadows View to assess how shadows are cast by trees and buildings. The specific observation behavior/action matters less than the fact that learners are making observations to develop an understanding of the task. Thus, all these events can be aggregated as observation behavior. Similar aggregation can occur for instances of formulation, analysis, reformulation, and evaluation. Such aggregated computer trace data can be used to examine the occurrence and frequency of SRL activities. Furthermore, advanced data mining techniques, such as sequential mining, can be used to reveal how SRL patterns are related to students' performance.

Sequential mining is becoming an increasingly valuable analytical tool for identifying behavioral patterns, especially in the investigation of SRL (Kinnebrew et al., 2013). Differential sequential mining algorithms have been used to compare behavioral patterns of high versus low efficient students (Zheng et al., 2022), high versus low performers (Kinnebrew et al., 2017), and students in experimental versus control conditions (Wong et al., 2019). For example, Kinnebrew et al. (2017) conducted a study investigating the patterns in strategy use between high and low performers. They found that sequential mining was an effective method in detecting students' learning strategy patterns as they engaged with Betty's Brain, an open-ended learning environment. Zheng et al. (2022) utilized sequential mining to examine students' SRL patterns in clinical reasoning, revealing that less efficient students exhibited more disorganized behavior than efficient students. In another study, Taub et al. (2018) used sequential pattern mining to determine if there were differences in hypothesis testing behaviors between high and low-efficient students when they played a game about microbiology on Crystal Island. They found that low-efficient students had the most sequences of testing behaviors, especially less strategic hypothesis testing behaviors. As demonstrated by these previous studies, sequential mining techniques unveil crucial behavioral patterns that significantly advance our understanding of the sequential and temporal aspects of SRL. Additionally, these studies serve as exemplary models for investigating differential behavioral patterns by employing sequential mining techniques in the context of the engineering design process.

3. Current study

The purpose of this study is to gain a better understanding of how computer trace data can be used to capture the temporal nature of SRL in engineering design. Using sequential mining and performance data during an engineering design task, we addressed the following research questions: (1) Do high-performing and low-performing students differ in the frequency of their enactment of SRL processes? (2) Do high-performing and low-performing students demonstrate different sequential patterns of self-regulatory processes? The first research question aims to uncover relevant differences between high and low-performing students in SRL. We anticipate that a significantly higher frequency of SRL sequences would be observed in the high-performing group compared to the low-performing group. The second research question delves into the temporal nature of SRL, specifically the difference in sequential patterns between high and low-performing students. Based on previous research findings (Zheng et al., 2020), we anticipate observing more consistent sequential patterns in the high-performing group compared to the low-performing group.

4. Methods

4.1. Participants

This is part of a larger study that involves students from a suburban high school in the northeastern United States. This study included 101 students who completed the learning task and consented to participate. Among them, fifty-five (50.9%) were female and fifty-three (49.1%) were male. These participants were enrolled in five honors courses in physical science, instructed by a male teacher with more than 17 years of experience teaching physical sciences and 5 years of experience guiding engineering design projects. Based on the school's data, the majority of the student population consists of White students, accounting for 76.7% of the total. The remaining percentages comprise Hispanic students at 4.6%, African American students at 4.2%, students of multiple races at 3.4%, Native American students at 0.2%, and Native Hawaiian/Pacific Islander students at 0.2%.

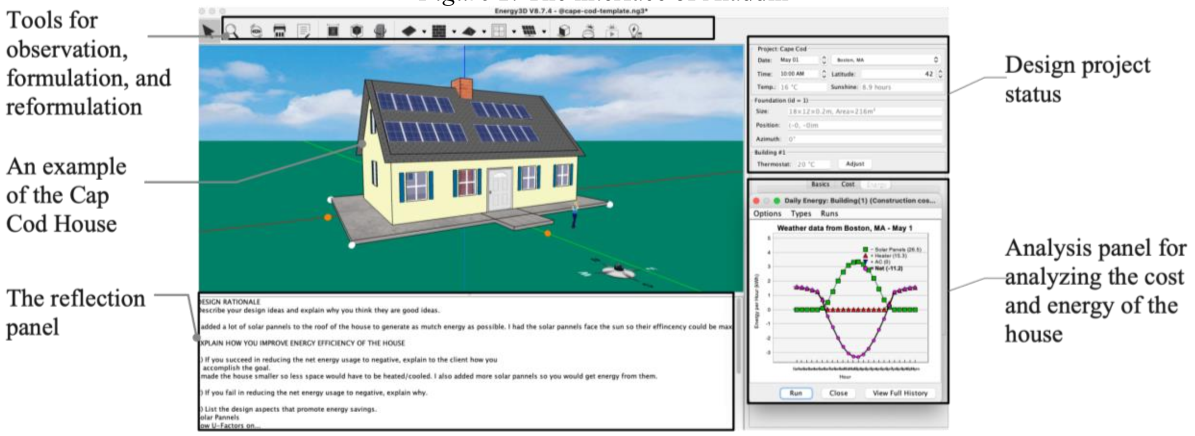
4.2. The learning environment

Aladdin, previously known as Energy3D, is a computer-supported simulation environment that supports engineering design, and it enables the construction, designing, and evaluation of green buildings that utilize solar power (Xie et al., 2018). Aladdin (see Figure 2) offers various simulations to aid students in understanding the solar conditions in which their buildings would be situated, including showing shadows, heliodon displays, and sun animations (i.e., observations). Figure 2 also demonstrates how Aladdin provides students with tools for 3D modeling, including walls, windows, roofs, solar panels, and trees. These tools allow for quick and easy creation and modification of building designs (i.e., formulation and reformulation). Additionally, Aladdin offers a wide range of quantitative analysis tools (such as energy and solar energy calculations, and annual energy analysis), which aid users in assessing the energy efficiency of their buildings and making necessary modifications to meet the net energy requirement. Finally, students are encouraged to reflect and evaluate their every round of design by responding to built-in prompts such as, "Describe your design ideas and explain why you think they are good ideas." Aladdin maintains a timeline that records all user actions, and these actions can reveal the temporal regulatory processing of engineering design.

4.3. Procedures

The institutional review board approved this study to protect the rights of human participants. All students who took part in this research provided their consent by filling out consent forms. The study was conducted in a science course, where participants were given 50-80 minutes each day to design a house using Aladdin software. The participants dedicated two days to the task of designing a single house. On the first day, participants got familiar with the Aladdin platform with the guidance of two researchers. They learned how to construct buildings, modify buildings, utilize the built-in simulations, and perform daily/annual energy analyses. At the same time, they learned the specified design requirements that were listed in a two-page printed instruction. Students are supposed to use Aladdin to individually design an energy-plus Cape Cod style house. To qualify as an energy-plus building, the house is required to generate a surplus of renewable energy over the course of 1 year, surpassing the amount of energy it utilizes. The Cape Cod style house is expected to fulfill certain criteria: it should exhibit an attractive exterior design; the ratio of the overall window area to the floor area should range between 0.05 and 0.15; trees should be positioned no closer than 2 meters from the house walls; the roof overhang cannot exceed a width of 50 centimeters. The allocated budget for the construction of the house is $200,000. The living space should range from 100 to 150 square meters, and the height of the house should be between 7 and 9 meters. In addition to design requirements, a two-page instructional handout also included an engineering design cycle to assist students in their design process. Students were encouraged to exchange ideas with their classmates before creating a building using Aladdin. On the second day, students started to perform construction independently. After finishing the construction, students were able to iterate their design by evaluating the energy efficiency of their building with embedded analysis tools. In total, students then spent two consecutive days completing the Cap Cord design of a house. Other than the provided instruction manual, students received very little explicit direction or guidance.

Figure 2. The interface of Aladdin



4.4. Data analysis

A total of 70,236 lines of student-generated computer trace data were utilized in this study. Each action conducted by students in the Aladdin environment was recorded by one row of data, including student anonymized ID, timestamped computer actions (by second), and action types. Each action was manually coded by researchers in alignment with the domain-specific SRL model (Figure 1) to understand students' self-regulatory phases, or processes. The coding scheme, as presented in Table 1, provided the operational definition for each SRL process along with the corresponding student actions. For instance, actions such as "show heliodon," "show shadow," and "animate sun" were used by students to monitor the movement of the sun throughout the day or year. Thus, these actions were coded as part of the Observation process within the engineering design domain.

Table 1. Coding scheme of SRL

Self-regulatory processes	Definition	Computer actions
Observation	Students observe the building and the sun using different types of visualization tools	Show Heliodon; Show Shadow; Animate Sun; Spin View; Top View; Show Annotation; Show Axes
Formulation	Students create the building by adding different components to the house	Add Solar Panel; Add Window; Add Wall; Add Trees; Add Hip Roof; Add Door; Add Pyramid Roof; Add Custom Roof; Add Floor; Add Foundation; Add people; Add Sensor;
Reformulation	Students modify the house by changing the size, location, and attributes of different house components and discarding some house components	Edit Window; Edit Solar Panel; Edit Wall; Change Date; Remove Solar Panel; Remove Window; Remove Trees; Remove Wall; Efficiency Change for Selected Solar Panel; Move Trees etc.
Analysis	Students calculate the energy production and consumption of the house	Solar Energy; Compute Energy; Energy Annual Analysis; Cost; Annual Sensor Data; Energy Angular Analysis
Evaluation	Students reflect on the design processes by justifying their decision-making process.	Make notes

The data was processed in Google Colab using Python programming language. First, we categorized students into high and low-performance groups based on their engineering design performance. Students' energy performance was assessed by calculating the net annual energy of the Cape Cod style house they constructed. The net annual energy of a house is determined by subtracting its annual production energy from its annual consumption energy. A negative value denotes an energy-plus house. A lower net energy value signifies a more energy-efficient house. The 101 participants were divided into a high-performing group (n = 51) and a low-performing group (n = 50) by using the 50th percentile as the threshold. To address the first research question, we presented the frequency of all five SRL processes to provide a holistic perspective on students' engagement in SRL phases during the engineering design process. In addition, the SRL processes were transformed into sequences that reflect students' SRL patterns and the iterative nature of the design process. Table 2 presents some examples of generated sequences. As displayed in Table 2, the numbers in the third column represent different SRL processes (Observation = 1, Formulation = 2, Reformulation = 3, Analysis = 4, Evaluation = 5). More specifically, the first row indicates that student A conducted a design activity consisting of five SRL processes on March 13th, while the same student performed a different sequential pattern the next day. In total, 310 SRL sequences were generated by all participants throughout the engineer design activity. Furthermore, we transformed the SRL sequences into a series of two-action sequences (see Table 3). As indicated in Table 3, the columns named "Sequences" present two-action combinations between process 1 and process 5, which resulted in 20 pairs. For instance, the cell "1-2" represents the sequential actions students conducted from the Observation (1) phase to the Formulation (2) conducted by students, which is 152 in this case. Similarly, for 1-3 under column "Sequences," the corresponding "Frequencies" is 304. This indicates that students conducted a total of 304 sequential actions from Observation (1) to Reformulation (3) throughout the entire activity.

Table 2. Example of SRL sequences

Student	Date	Activity sequence
A	Mar.13th	[1, 5, 2, 3, 4]
A	Mar.14th	[5, 3, 4, 5, 4, 5, 5, 3, 5, 4, 5]
C	Mar.17th	[3, 4, 1, 2, 3, 4, 3, 4, 3, 4, 5, 2, 3, 2]

Note. 1 = Observation, 2 = Formulation, 3 = Reformulation, 4 = Analysis, 5 = Evaluation.

Table 3. The frequencies of consecutive activity pairs

Sequences	Frequencies
1-2	152
1-3	304
1-4	181
1-5	30
2-1	99
2-3	2399
2-4	545
2-5	23
3-1	291
3-2	2322
3-4	1982
3-5	111
4-1	208
4-2	461
4-3	1851
4-5	335
5-1	77
5-2	147
5-3	222
5-4	235

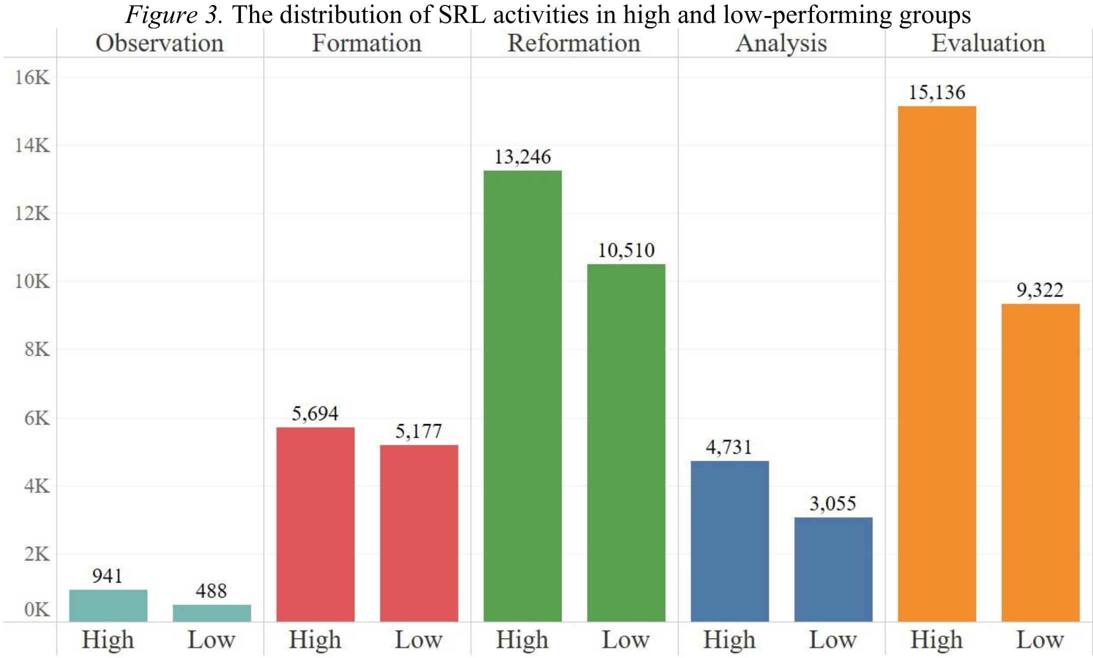
Note. 1 = Observation, 2 = Formulation, 3 = Reformulation, 4 = Analysis, 5 = Evaluation.

5. Results

5.1. Research Question 1. Do high-performing and low-performing students differ in the frequency of their enactment of SRL processes?

To answer the first research question, participants' behavioral frequencies that mapped into correspondent SRL phases were divided into high (n = 51) and low-performing (n = 50) groups. As displayed in Figure 3, reformulation and evaluation were the most frequent SRL activities in both high and low-performing groups, indicating that students devoted the most effort to revising the design and reflecting on the design processes. In general, high-performing students performed more SRL activities in all phases.

Figure 3. The distribution of SRL activities in high and low-performing groups



In particular, independent t-tests were performed on the frequencies of five SRL phases on students from high and low-performing groups. As indicated in Table 4, there were significant differences between the two groups in observation, analysis, and evaluation. More specifically, in the observation phase (t (82) = 2.32, p < .05), students in the high-performing group (M = 21.39, SD = 21.76) presented significantly higher frequencies than the low-performing group (M = 12.20, SD = 12.77) with a medium effect size (d = .51). Students in the high-performing group (M = 92.76, SD = 80.81) also conducted significantly more analysis actions ((t (99) = 2.60, p < .05) than that of low-performing group (M = 61.10, SD = 59.25) with a moderate effect size (d = .45). Regarding the Evaluation phase, students in the high-performing group (M = 296.78, SD = 199.19) exhibited significantly higher frequencies compared to the low-performing group (M = 186.44, SD = 220.71) with a medium effect size (d = .53).

Table 4. Comparison of SRL activities frequency between high and low performing group

SRL phase	Group	*N*	*M*	*SD*	*t*	*df*	*p*	Cohen's *d*
Observation	High	44	21.39	21.76	2.32	82	.02*	.51
	Low	40	12.20	12.77				
Formation	High	51	111.65	58.46	.64	99	.52	.13
	Low	50	103.54	68.72				
Reformation	High	51	259.73	156.70	1.68	99	.10	.34
	Low	50	210.20	138.28				
Analysis	High	51	92.76	80.81	2.25	99	.03*	.45
	Low	50	61.10	59.25				
Evaluation	High	51	296.78	199.19	2.60	99	.01*	.53
	Low	50	186.44	220.71				

Note. $^{*}p < .05$.

5.2. Research Question 2. Do high-performing and low-performing students demonstrate different sequential patterns of self-regulatory processes?

Students in the high-performing group (n = 51) performed 40,798 log actions that equivalented to 179 SRL sequences, whereas the low-performing group (n = 50) conducted 29,438 log actions that equivalented to 131 SRL sequences. Sankey graphs were generated to visualize students' sequential and iterative SRL processes in engineering design. A Sankey graph is a particular type of data visualization that employs arrows of varying widths to depict the flow of quantities between different phases (Pan & Liu, 2022). As displayed in Figure 4 and Figure 5, high and low-performing students demonstrated similar iterative SRL patterns. Students focus the majority of their effort on the iteration of Formulation, Reformulation, and Analysis, suggesting that these three actions play a significant role in determining the flow of the engineering design process. Although the iterative SRL patterns were similar, we conducted further exploration to determine if there were differences in the sequential pattern and directionality of the sequential pattern between high and low-performing students.

Figure 4. The sequential patterns of high-performing students' regulatory processes

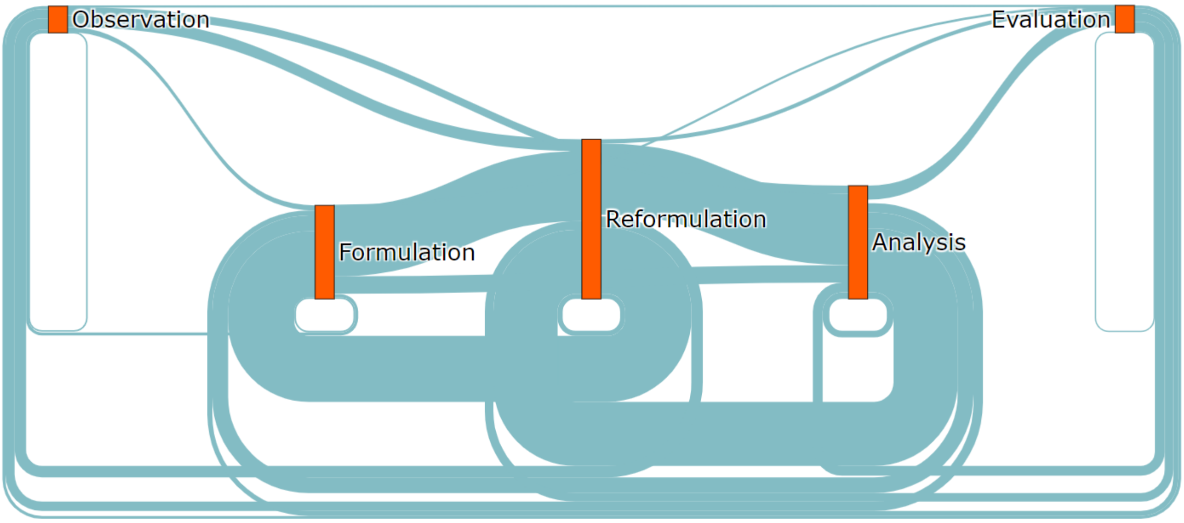
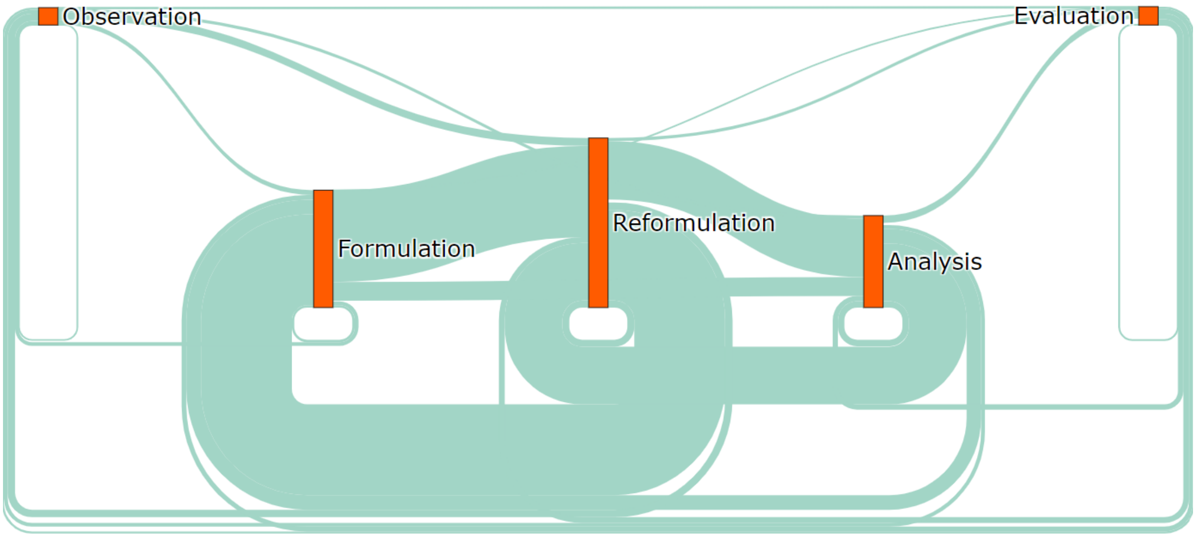


Figure 5. The sequential patterns of low-performing students' regulatory processes



Non-parametric Mann-Whitney U tests were conducted to compare the difference in the frequency of SRL sequence between the two groups since the data violates the normality assumption. Table 5 presents the results generated by Mann-Whitney U tests. More specifically, it presents the average and total frequency of each SRL sequence for both high and low-performing groups. Additionally, the table includes the Mann-Whitney U value and p value to signify any significant difference between two groups in each SRL sequence. For example, for the sequence 1-4 (Observation to Analysis), students in the high-performing group performed this sequential action a total of 139 times. On average, each student conducted this sequential action 2.73 times. In contrast, students in the low-performing group only conducted this action 42 times in total, with an average of 0.88 times per student. This sequential action showed a significant difference between the two groups ($U = 886.50$, $p < .05$).

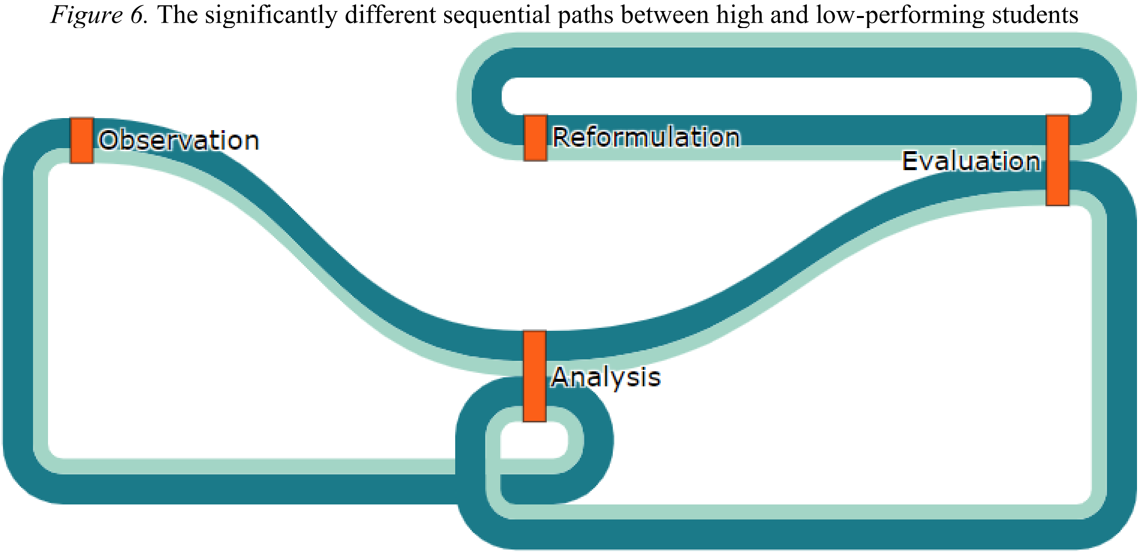
As indicated in Table 5 and Figure 6, there were significant differences in six different SRL sequences or three SRL cynical patterns between the two groups: Observation <=> Analysis (1-4, 4-1); Reformulation <=> Evaluation (3-5, 5-3); Analysis <=> Evaluation (4-5, 5-4). More specifically, students in the high-performance group demonstrated significantly more sequential patterns from Observation to Analysis phase (1-4, $U = 886.50$, $p < .05$, $r = .28$) and from Analysis to Observation (4-1, $U = 921.00$, $p < .05$, $r = .25$) than their peers in the low-performance group. For the cynical pattern between Reformulation and Evaluation, the high-performing group also conducted significantly more than the low-performing group in the sequence from Reformulation to Evaluation (3-5, $U = 963.50$, $p < .05$, $r = .23$) and from Evaluation to Reformulation (5-3, $U = 826.50$, $p < .05$, $r = .31$). Moreover, the high-performance group also conducted significantly more sequences from Analysis to Evaluation (4-5, $U = 628.50$, $p < .05$, $r = .44$) and from Evaluation to Analysis phase (5-4, $U = 683.50$, $p < .05$, $r = .41$) than the low-performing group.

Table 5. Mann-Whitney U tests results comparing the sequential patterns of high and low-performing students

Sequence	High performance group		Low performing group		Mann-Whitney U
	Mean	Total	Mean	Total	
1-2	1.73	88	1.28	64	1196.00
1-3	3.94	201	2.06	103	1004.50
1-4	**2.73***	**139**	**0.84**	**42**	**886.50**
1-5	0.29	15	0.30	15	1154.00
2-1	0.98	50	0.98	49	1306.50
2-3	23.33	1190	24.18	1209	1329.50
2-4	5.80	296	4.98	249	1187.50
2-5	0.35	18	0.10	5	1027.50
3-1	3.78	193	1.96	98	1033.50
3-2	21.98	1121	24.02	1201	1407.00
3-4	23.73	1210	15.44	772	1008.00
3-5	**1.37***	**70**	**0.82**	**41**	**963.50**
4-1	**3.10***	**158**	**1.00**	**50**	**921.00**
4-2	5.25	268	3.86	193	1087.50
4-3	21.33	1088	15.26	763	1097.50
4-5	**4.69***	**239**	**1.92**	**96**	**628.50**
5-1	0.96	49	0.56	28	1155.50
5-2	1.65	84	1.26	63	1032.50
5-3	**2.86***	**146**	**1.52**	**76**	**826.50**
5-4	**3.25***	**166**	**1.38**	**69**	**683.50**

Note. $^{*}p < .05$. 1 = Observation, 2 = Formulation, 3 = Reformulation, 4 = Analysis, 5 = Evaluation.

Figure 6. The significantly different sequential paths between high and low-performing students



6. Discussion

This study found that students allocated the majority of their effort to evaluation and reformulation processes, signifying a strong focus on the self-reflection and performance phase of SRL. The significant allocation of effort to these two phases highlights the active involvement and metacognitive awareness of students in evaluating and revising their designs. Although there are no prior findings on how students distribute their effort across the three phases of SRL, it is important to note that effective self-regulated learners do not evenly distribute their effort across the three phases of SRL even though all three phases play important roles in SRL (Li et al., 2018). The three phases of SRL are interlinked, collectively shaping the learning process. For example, Callan and Cleary (2019) explored the connection between SRL phases, revealing that effort invested in the forethought phase positively influenced effort in the performance phase, subsequently impacting mathematics performance. However, an overemphasis on the forethought phase might lead to reduced time for the performance phase, potentially resulting in poor academic performance (Li et al., 2018). Similarly, learners who do not engage in self-reflection may miss opportunities to adjust their learning strategies to improve learning efficiency and performance. Therefore, it is essential for learners to balance their effort across the three phases of SRL, especially when heavily involved in the self-reflection and performance phase.

Additionally, we found that high-performing students devoted more effort than low-performing students to the three phases of SRL, particularly the Observation, Analysis, and Evaluation processes. This finding is consistent with prior studies that contended that high-performing students devote more time to all phases of SRL compared to lower performers (Dibenedetto & Zimmerman, 2010; DiFrancesca et al., 2016; Foong et al., 2021; Li et al., 2018). This finding also supports our previous argument that learners should balance their effort across the phases of SRL. Moreover, in the context of engineering design, a distinctive finding emerges where high-performing students show deeper engagement in analyzing design solutions during the performance phase of SRL. This finding highlights that high-performing students allocate considerable effort to critically analyze the effectiveness of their design solutions. This analytical strategy enables them to identify areas of strength and areas that require improvement (Katz, 2015), leading to adjustments that optimize their learning process. Thus, educators can use this finding to design targeted interventions and provide support to students who may need assistance in analytic thinking skills in the performance phase.

Of particular interest, students in the high-performing group exhibited a higher occurrence of consecutive sequential patterns between Observation and Analysis, Reformation and Evaluation, and Analysis and Evaluation. The frequent occurrence of consecutive patterns between Observation and Analysis shows that high-performing students systematically gather relevant information and then analyze it effectively to make informed decisions about their design solutions. The high incidence of consecutive patterns between Reformation and Evaluation suggests that high-performing students actively evaluate their design after implementing changes or reforms in their learning strategies. The frequent consecutive patterns between Analysis and Evaluation imply that high-performing students adeptly utilize their analytical skills to assess their performance and learning outcomes. These three sequential patterns indicate that high-performing students possess metacognitive awareness (Ramirez-Corona et al., 2013), actively engaging in analysis based on observation, evaluating reformed design solutions, and integrating analytical thinking into their evaluation process. These self-assessment and self-adjustment processes enable students to refine their design solutions and optimize their design performance. Educators and computer program developers can use these findings to promote metacognitive development in engineering design, highlighting the importance of observational skills, critical analysis, and self-evaluation throughout the learning journey. By designing prompts that reinforce the connection between observation, analysis, and evaluation, educators can encourage students to engage more actively in their self-assessment and self-adjustment processes.

7. Limitations and future directions

Fine-grained tracing of student actions and sequential mining of their temporal patterns offer insights into the temporal structures of students' SRL in engineering design. However, this study also comes with limitations. The computer trace data was collected in a single task (i.e., designing a Cape Cod Style energy-plus house) and thus generalizability is constrained to that task and sample. As previously discussed, SRL strategies and behaviors employed by individuals can vary based on the nature of the learning task (Alexander et al., 2011). Consequently, the identified sequential patterns may vary when considering different contexts and individuals. Future research could replicate this study across various engineering design tasks to validate the findings. Moreover, results may be further influenced by external support during the task, such as peer support and teacher support. Although peer support and teacher support were not available in the current study, this kind of external support may influence the efficiency and efficacy of SRL. Thus, future research should take into account these external factors when analyzing students' self-regulatory processes. Another limitation pertains to the method of categorizing students into high and low-performing groups. While design outcomes hold significant importance in this task, they are not the sole metrics for differentiation. For instance, the aesthetic dimension of the design could be considered in evaluating students' performance. Aspects such as students' motivation and prior knowledge also matter. Future studies should consider measuring students' motivation, prior knowledge, and the aesthetic dimension of the design to differentiate students comprehensively. Finally, the study did not establish causal relationships despite the fact that we successfully examined the differences in SRL patterns between high and low-performing students. To further validate the efficacy of these findings, future research endeavors could delve into examining the impact of interventions crafted based on these identified patterns.

8. Conclusion

This study examined the SRL patterns of 101 high school students as they performed engineering design in a computer-supported simulation environment, Aladdin. The sequential mining results revealed that students in the high-performing group engaged in significantly more actions associated with the Observation, Analysis, and Evaluation processes of SRL. Moreover, students from high and low-performing groups exhibited distinct self-regulatory patterns. The visualized sequential patterns, presented through Sankey charts, depicted more nuanced differences between the two groups, including iterative SRL patterns and variations in the proportions of the paths taken by each group. More specifically, students in the high-performing group presented a higher occurrence of consecutive sequential patterns between Observation and Analysis, between Reformation and Evaluation, and between Analysis and Evaluation.

The results have practical implications for the design of interventions and system development. First, educators can tailor interventions to assist students in honing their analytical thinking skills during the performance phase. Secondly, educators and computer program developers can collaborate to foster metacognitive growth in engineering design by emphasizing observational skills, critical analysis, and self-evaluation throughout the learning process. By incorporating prompts that reinforce the interconnectedness of observation, analysis, and evaluation, educators can stimulate students' active engagement in self-assessment and adjustment. Lastly, these results underscore the importance of evenly distributing effort across the three phases of SRL for achieving effective learning outcomes.

Acknowledgement

This project is supported by the National Science Foundation (NSF) under grant numbers #2105695, #2131097, and #2301164. Any opinions, findings, and conclusions or recommendations expressed in this material, however, are those of the authors and do not necessarily reflect the views of NSF.

References

Alexander, P. A., Dinsmore, D. L., Parkinson, M. M., & Winters, F. I. (2011). Self-regulated learning in academic domains. In D. Schunk & B. Zimmerman (Eds.), *Handbook of self-regulation of learning and performance* (1st ed., pp. 393–407). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780203839010

Apedoe, X. S., Reynolds, B., Ellefson, M. R., & Schunn, C. D. (2008). Bringing engineering design into high school science classrooms: The heating/cooling unit. *Journal of Science Education and Technology*, *17*(5), 454–465. https://doi.org/10.1007/s10956-008-9114-6

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 427–449). Springer New York. https://doi.org/10.1007/978-1-4419-5546-3_28

Azevedo, R., Taub, M., & Mudrick, N. V. (2018). Understanding and reasoning about real-time cognitive, affective, and metacognitive processes to foster self-regulation with advanced learning technologies. In D. Schunk & J. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 254–270). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315697048-17

Barroso, L. R., Nite, S. B., Morgan, J. R., Bicer, A., Capraro, R. M., & Capraro, M. M. (2016). Using the engineering design process as the structure for project-based learning: An informal STEM activity on bridge-building. In *Proceedings of 2016 IEEE Integrated STEM Education Conference (ISEC)* (pp. 249–256). https://doi.org/10.1109/ISECon.2016.7457542

Becker, K., & Mentzer, N. (2015). Engineering design thinking: High school students' performance and knowledge. *2015 International Conference on Interactive Collaborative Learning (ICL)*, 5–12. https://doi.org/10.1109/ICL.2015.7318218

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), 1–13. https://doi.org/10.1007/s11409-015-9134-6

Beneroso, D., & Robinson, J. (2022). Online project-based learning in engineering design: Supporting the acquisition of design skills. *Education for Chemical Engineers*, *38*, 38–47. https://doi.org/10.1016/j.ece.2021.09.002

Bernacki, M. L. (2017). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In D. Schunk & J. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 370–387). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315697048-24

Callan, G. L., & Cleary, T. J. (2019). Examining cyclical phase relations and predictive influences of self-regulated learning processes on mathematics task performance. *Metacognition and Learning*, *14*(1), 43–63. https://doi.org/10.1007/s11409-019-09191-x

Das, M., & Yang, M. C. (2021). Design experiences as pathways for embracing failure. In *Proceedings of ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference.* https://doi.org/10.1115/DETC2021-71419

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8

Dibenedetto, M., & Zimmerman, B. (2010). Differences in self-regulatory processes among students studying science: A microanalytic investigation. *International Journal of Educational & Psychological Assessment*, *5*, 2–24.

DiFrancesca, D., Nietfeld, J. L., & Cao, L. (2016). A comparison of high and low achieving students on self-regulated learning variables. *Learning and Individual Differences*, *45*, 228–236. https://doi.org/10.1016/j.lindif.2015.11.010

Dym, C. L. (2013). *Engineering design: A project-based introduction*. John Wiley & Sons.

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE—Life Sciences Education*, *13*(3), 453–468. https://doi.org/10.1187/cbe.14-03-0050

Fan, Y., Tan, Y., Raković, M., Wang, Y., Cai, Z., Shaffer, D. W., & Gašević, D. (2023). Dissecting learning tactics in MOOC using ordered network analysis. *Journal of Computer Assisted Learning*, *39*(1), 154–166. https://doi.org/10.1111/jcal.12735

Foong, C. C., Bashir Ghouse, N. L., Lye, A. J., Khairul Anhar Holder, N. A., Pallath, V., Hong, W.-H., Sim, J. H., & Vadivelu, J. (2021). A qualitative study on self-regulated learning among high performing medical students. *BMC Medical Education*, *21*, 320. https://doi.org/10.1186/s12909-021-02712-w

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Greene, J. A., Plumley, R. D., Urban, C. J., Bernacki, M. L., Gates, K. M., Hogan, K. A., Demetriou, C., & Panter, A. T. (2021). Modeling temporal self-regulatory processing in a higher education biology course. *Learning and Instruction*, *72*, 101201. https://doi.org/10.1016/j.learninstruc.2019.04.002

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in Introductory Biology. *Science*, *332*(6034), 1213–1216. https://doi.org/10.1126/science.1204820

Karan, E., & Brown, L. (2022). Enhancing students' problem-solving skills through project-based learning. *Journal of Problem Based Learning in Higher Education*, *10*(1), 74-87. https://doi.org/10.54337/ojs.jpblhe.v10i1.6887

Katz, R. (2015). Integrating analysis and design in mechanical engineering education. *Procedia CIRP*, *36*, 23–28. https://doi.org/10.1016/j.procir.2015.01.042

Kennedy, T. J., & Odell, M. R. L. (2014). Engaging students in STEM education. *Science Education International*, *25*(3), 246–258.

Kinnebrew, J. S., Loretz, K. M., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining*, *5*(1), 190–219. https://doi.org/10.5281/zenodo.3554617

Kinnebrew, J. S., Segedy, J. R., & Biswas, G. (2017). Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments. *IEEE Transactions on Learning Technologies*, *10*(2), 140–153. https://doi.org/10.1109/TLT.2015.2513387

Kober, N. (2015). *Reaching students: What research says about effective instruction in undergraduate science and engineering*. National Academies Press. https://doi.org/10.17226/18687

Kuppuswamy, R., & Mhakure, D. (2020). Project-based learning in an engineering-design course – Developing mechanical-engineering graduates for the world of work. *Procedia CIRP*, *91*, 565–570. https://doi.org/10.1016/j.procir.2020.02.215

Lawanto, O., Butler, D., Cartier, S. C., Santoso, H. B., Goodridge, W., Lawanto, K. N., & Clark, D. (2013 a). Pattern of task interpretation and self-regulated learning strategies of high school students and college freshmen during an engineering design project. *Journal of STEM Education: Innovations and Research*, *14*(4), 15-27.

Lawanto, O., Butler, D., Cartier, S., Santoso, H., Lawanto, K., & Clark, D. (2013 b). An exploratory study of self-regulated learning strategies in a design project by students in grades 9-12. *Design and Technology Education*, *18*(1), 44–57.

Li, S., Zheng, J., Poitras, E., & Lajoie, S. (2018). The allocation of time matters to students' performance in clinical reasoning. In R. Nkambou, R. Azevedo, & J. Vassileva (Eds.), *Intelligent Tutoring Systems* (pp. 110–119). Springer International Publishing. https://doi.org/10.1007/978-3-319-91464-0_11

Lin, C.-H., Yu, C.-C., Shih, P.-K., & Wu, L. Y. (2021). STEM based artificial intelligence learning in general education for non-engineering undergraduate students. *Educational Technology & Society*, *24*(3), 224–237.

Lin, K.-Y., Hsiao, H.-S., Chang, Y.-S., Chien, Y.-H., & Wu, Y.-T. (2018). The effectiveness of using 3D printing technology in STEM project-based learning activities. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(12), em1633. https://doi.org/10.29333/ejmste/97189

Lönngren, J., Adawi, T., Berge, M., Huff, J., Murzi, H., Direito, I., Tormey, R., & Sultan, U. (2020). Emotions in engineering education: Towards a research agenda. *2020 IEEE Frontiers in Education Conference (FIE)*, 1–5. https://doi.org/10.1109/FIE44824.2020.9273951

McFadden, J., & Roehrig, G. (2019). Engineering design in the elementary science classroom: Supporting student discourse during an engineering design challenge. *International Journal of Technology and Design Education*, *29*(2), 231–262. https://doi.org/10.1007/s10798-018-9444-5

Mentzer, N. (2014). High school student information access and engineering design performance. *Journal of Pre-College Engineering Education Research (J-PEER)*, *4*(1). https://doi.org/10.7771/2157-9288.1074

Mentzer, N., Becker, K., & Sutton, M. (2015). Engineering design thinking: High school students' performance and knowledge. *Journal of Engineering Education*, *104*(4), 417–432. https://doi.org/10.1002/jee.20105

Palid, O., Cashdollar, S., Deangelo, S., Chu, C., & Bates, M. (2023). Inclusion in practice: A systematic review of diversity-focused STEM programming in the United States. *International Journal of STEM Education*, *10*(1), 2. https://doi.org/10.1186/s40594-022-00387-3

Pan, Z. & Liu, M. (2022). Theory-Informed Problem-Solving Sequential Pattern Visualization. In *Proceedings of the 15th International Conference on Educational Data Mining (EDM'22)* (pp. 738-742). International Educational Data Mining Society (EDM).

Poitras, E. G., & Lajoie, S. P. (2013). A domain-specific account of self-regulated learning: The cognitive and metacognitive activities involved in learning through historical inquiry. *Metacognition and Learning*, *8*(3), 213–234. https://doi.org/10.1007/s11409-013-9104-9

Ramirez-Corona, N., Zaira, R. A. L., Lopez-Malo, A., & Palou, E. (2013, June 23-26). *Assessing metacognitive awareness during problem-solving in a kinetics and homogeneous reactor design course* [Conference Presentation]. 2013 ASEE Annual Conference & Exposition, Atlanta, Georgia. https://doi.org/10.18260/1-2--19227

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, *36*(1), 111–139. https://doi.org/10.1007/s11165-005-3917-8

Sinatra, G. M., & Taasoobshirazi, G. (2017). The self-regulation of learning and conceptual change in science: Research, theory, and educational applications. In D. Schunk & J. Greene (Eds.), *Handbook of Self-Regulation of Learning and Performance* (2nd ed., pp. 153-165). Routledge. https://doi.org/10.4324/9781315697048-10

Taub, M., Azevedo, R., Bradbury, A. E., Millar, G. C., & Lester, J. (2018). Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and Instruction*, *54*, 93–103. https://doi.org/10.1016/j.learninstruc.2017.08.005

Wong, J., Khalil, M., Baars, M., de Koning, B. B., & Paas, F. (2019). Exploring sequences of learner activities in relation to self-regulated learning in a massive open online course. *Computers & Education*, *140*, 103595. https://doi.org/10.1016/j.compedu.2019.103595

Xie, C., Schimpf, C., Chao, J., Nourian, S., & Massicotte, J. (2018). Learning and teaching engineering design through modeling and simulation on a CAD platform. *Computer Applications in Engineering Education*, *26*(4), 824–840. https://doi.org/10.1002/cae.21920

Zheng, J., Xing, W., Huang, X., Li, S., Chen, G., & Xie, C. (2023). The role of self-regulated learning on science and design knowledge gains in engineering projects. *Interactive Learning Environments*, *31*(1), 87-99. https://doi.org/10.1080/10494820.2020.1761837

Zheng, J., Xing, W., Zhu, G., Chen, G., Zhao, H., & Xie, C. (2020). Profiling self-regulation behaviors in STEM learning of engineering design. *Computers & Education*, *143*, 103669. https://doi.org/10.1016/j.compedu.2019.103669

Zheng, J., Li, S., & Lajoie, S. P. (2022). Diagnosing virtual patients in a technology-rich learning environment: A sequential mining of students' efficiency and behavioral patterns. *Education and Information Technologies*, *27*(3), 4259–4275. https://doi.org/10.1007/s10639-021-10772-0

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, *48*(3), 135–147. https://doi.org/10.1080/00461520.2013.794676