Computers & Education 143 (2020) 103669

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Profiling self-regulation behaviors in STEM learning of engineering design

Juan Zheng[a,e], Wanli Xing[b,e,*], Gaoxia Zhu[c,e], Guanhua Chen[d,e], Henglv Zhao[c,e], Charles Xie[d,e]

[a] Department of Educational Counselling & Psychology, McGill University, Montreal, QC, H3A 1Y9, Canada
[b] School of Teaching & Learning, University of Florida, Gainesville, FL 32601, USA
[c] Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, M5S 1V6, Canada
[d] The Concord Consortium, Concord, MA, 01742, USA
[e] Learning Genome Collaborative, Natick, MA, USA

ARTICLE INFO

Keywords:
Self-regulation
STEM learning
Learning analytics
Energy3D
Engineering design

ABSTRACT

Engineering design is a complex process which requires science, technology, engineering, and mathematic (STEM) knowledge. Students' self-regulation plays a critical role in interdisciplinary tasks. However, there is limited research investigating whether and how self-regulation leads to different learning outcomes among students in engineering design. This study analyzes the engineering design behaviors of 108 ninth-grade U.S. students using principal component analysis and cluster analysis. It classifies the students into four distinct types: competent, cognitive-oriented, reflective-oriented, and minimally self-regulated learners. Competent self-regulated learners perceived themselves as the most self-regulated learners and had the greatest learning gains, although they did not perform best in the task. Cognitive-oriented self-regulated learners perceived themselves as the least self-regulated learners although they were the second best in both the performance of the task and learning gains. In contrast, reflective learners had the best performance in the task. Minimally self-regulated learners did not perform well in the task and had the lowest learning gains. The results revealed that the competent self-regulated learners had an appropriate assessment of themselves to obtain knowledge, cognitive-oriented self-regulated learners underestimated themselves, reflective learners focused on the results of the task, and minimally self-regulated learners overestimated themselves and exerted the least effort. The results also offer new insights into STEM education and self-regulated learning with emerging learning analytics.

1. Introduction

There is an increasing call to facilitate STEM education through constructing integrated learning environments in which students can solve complex projects (Chiu et al., 2013). Engineering design activities usually involve engineering construction, science inquiry, mathematical reasoning, artistic design, and technological skills (Dasgupta, Magana, & Vieira, 2019), which can be considered as cross-discipline projects. With the infusion of engineering design into K-12 classrooms, educators recommend situating engineering problems by "posing authentic problems" and giving students access to "authentic practice" (Strobel, Wang, Weber, & Dyehouse,

* Corresponding author. School of Teaching & Learning, University of Florida, Gainesville, FL 32601, USA.
E-mail addresses: juan.zheng@mail.mcgill.ca (J. Zheng), wanli.xing@coe.ufl.edu (W. Xing), gaoxia.zhu@mail.utoronto.ca (G. Zhu), gchen@concord.org (G. Chen), Henglv.zhao@ttu.edu (H. Zhao), qxie@concord.org (C. Xie).

https://doi.org/10.1016/j.compedu.2019.103669
Received 24 April 2019; Received in revised form 28 July 2019; Accepted 17 August 2019
Available online 19 August 2019
0360-1315/ Published by Elsevier Ltd.

2013). The authenticity of problems and practice can be supported by computer-based learning environments where various simulation tools are available for unlimited practice. In addition to authenticity, computer-based environments make engineering design processes visible for educators and researchers towards better instructional practice.

Engineering design consists of dynamic learning processes where designers create models, test ideas, analyze data, and construct new knowledge to optimize design solutions (Crismond & Adams, 2012; Lewis, 2006). An engineering design activity usually involves high-order skills such as observing, modeling, modifying, analyzing, and evaluating a project (Fan & Yu, 2017), all of which draw on students' self-regulated learning (SRL).

SRL refers to the learning process where students actively monitor and control their learning using a variety of cognitive and behavioral strategies (Zimmerman, 2000). Indeed, past research shows that self-regulation plays an important role in students' efficiency and performance while completing an engineering design project (Lawanto & Johnson, 2012). For example, expert engineering designers were found to be more engaged in evaluating their design than novice designers (Dixon, 2010). Accordingly, it is very important to investigate if and how self-regulation may influence engineering design processes and consequently affect learning outcomes.

Although a wealth of research has demonstrated the impact of self-regulation on learning performance (Sitzmann & Ely, 2011), these studies have generally employed variable-oriented statistical approaches. These studies confirmed that SRL strategies are positively correlated with learning outcomes measured using students' self-reports (Buric & Soric, 2012; Ozcan, 2016; Peng, Hong, & Mason, 2014). However, how learners use self-regulated strategies and why different SRL processes may lead to different learning outcomes have not been fully explored. In addition, few studies have used behavioral-oriented trace data to determine students' self-regulation profiles. This is especially true for engineering design, which typically uses computer-based learning environments to support SRL. Self-regulated cognitive processes (e.g., evaluation) interact with the changes of mental representation in the course of navigating and solving an engineering design problem (Dixon, 2010). Furthermore, self-regulated learners can be distinguished by their awareness of the relations between SRL strategies and regulatory processes in their own processes (Zimmerman, 1990). Therefore, the purpose of this study is to 1) gain deeper insight into SRL in the field of STEM education by using a person-oriented approach to classify learners according to their engineering design behavior and 2) to examine whether and how learners with different self-regulated behavioral profiles differ in perceived engagement in metacognitive self-regulation and learning outcomes. The significance of this study is twofold: (1) different student SRL profiles during engineering design were identified when analyzing student design behaviors and the SRL profiles were found to be associated with their learning gains and design performance; and (2) the results may contribute to research and practice of using analytical tools to support students with different SRL profiles in the STEM design field, for example, providing minimally self-regulated learners who tend to overestimate themselves with prompts to regulate their design processes.

2. Conceptual framework and literature review

2.1. Self-regulated learning theory

SRL is operationalized by Zimmerman (1990, 2008, 2013) as dynamic and cyclical processes that consist of three independent phases: *forethought*, *performance*, and *self-reflection*. In the *forethought* phase, learners prepare efforts to learn by analyzing the task and establishing a specific goal for the task. The goal developed in this phase will guide the subsequent actions and behaviors of learners. In the *performance* phase, learners execute strategies to control and monitor their cognitive processes, which result in the progress or stagnation of the task. Finally, learners evaluate and optimize personal reactions to learning outcomes in the *self-reflection* phase. *Self-reflection* occurs when students receive internal or external feedback about the task. *Self-reflection* can trigger both momentum and obstruction for further efforts in SRL. For example, learners who evaluate themselves as insufficiently performed in the task can either react positively by committing more efforts to gain better learning outcomes or react negatively by diminishing their motivation for the task and even motivation for learning. This also explains the iterative and cyclical nature of SRL processes. The self-reflection in the first round may activate the start of further task analysis at the *forethought* phase or new strategies at *performance* phase in the second round. Zimmerman's (1990, 2008, 2013) model describes how SRL happens within a task at a general level. Serving as the primary theoretical and conceptual foundation of SRL, this model paves the way for understanding SRL in a specific domain or task (Cleary & Callan, 2018). This highly practical and explicit model can be adapted and expanded to study domain-specific regulatory processes that emerge from a specific learning activity.

Studying SRL at a general level is not enough to reveal the different regulatory processes associated with the different contents, tools, and strategies involved in a task. Domains (e.g., linguistics and engineering) or tasks (e.g. a reading task and a problem solving task) differ in terms of the nature of the subject and structure of the task, which influences the processes students may experience and strategies they may adopt to regulate their own learning (Poitras & Lajoie, 2013). For example, students may employ the strategies of rehearsal, selecting main ideas, and rewriting notes to prepare for a memory task (Dabbagh & Kitsantas, 2013), while they may favor systematic questioning and uncertainty reasoning when completing an engineering design task (Dym, Agogino, Eris, Frey, & Leifer, 2005). Very few empirical studies have examined this aspect. This is why Alexander, Dinsmore, Parkinson, and Winters (2011) called for further examination of SRL in different domains.

2.2. Self-regulated learning in engineering design

In the context of STEM education, engineering design requires particular knowledge schema and design processes. First of all,

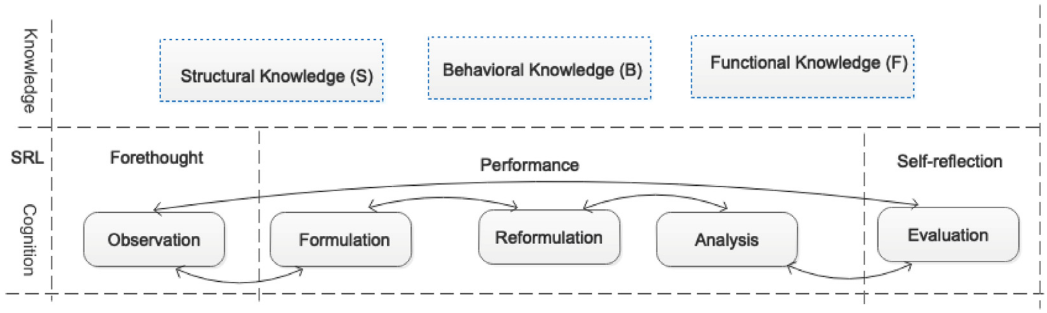

Fig. 1. SRL model in engineering design.

based on the design prototypes and design knowledge representation schema, researchers have generally conceived of engineering design projects as requiring *structural knowledge*, *functional knowledge*, and *behavioral knowledge* (Gero, 1990; Gero & Kannengiesser, 2004). *Structural knowledge* represents the components of a design project, such as walls and floors, in the project to design a house. *Functional knowledge* describes the technological purpose of the project. For example, the engineering design task in this study is to design a house that produces more renewable energy than it consumes over a year. *Behavioral knowledge* refers to the attributes that derive from the *structural knowledge*, such as the fact that windows can distribute the energy conserved in the house. *Behavioral knowledge* is usually implicit and functions as the connection between explicit *structural knowledge* and explicit *functional knowledge* (Gero & Kannengiesser, 2004). A good designer should appropriately combine the three types of knowledge. Secondly, *formulation, analysis, reformulation*, and *evaluation* are the well-recognized design processes in the literature (Howard, Culley, & Dekoninck, 2008). From the start to the end of a design project, designers transform their functional and structural knowledge into the components of the project (*formulation*), derive functional information from the structure (*analysis*), change and modify the project to conform to the intended design (*reformulation*), and compare the current structure and function with the intended design to assess if the design solution is acceptable (*evaluation*). In summary, knowledge schema and design processes depict the unique characteristics of the engineering design domain, which should be taken into consideration when developing a model of self-regulated learning in the domain of engineering design.

Referencing the existing models of SRL in a basic science field (Lajoie, Poitras, Doleck, & Jarrell, 2015) and taking into account the aforementioned knowledge schema and processes in engineering design, we developed a SRL model in engineering design (see Fig. 1). Since behavioral knowledge is implicit and the computer-supported learning environment we used for this study does not make it explicit, we focus on how learners use structural and functional knowledge to get through the SRL processes (i.e., forethought, performance, self-reflection). As displayed in Fig. 1, learners regulate themselves through five cognitive processes: *observation*, *formulation*, *reformulation*, *analysis*, and *evaluation*. Specifically, learners make *observations* to understand the task at the forethought phase. The performance phase is the critical phase where learners pursue the design task by *formulation*, *reformulation,* and *analysis*. Finally, learners *evaluate* if their current design matches the intended design in the self-reflection phase. All five of these cognitive processes are iterative and cyclical. For example, in the context of this study, students who are supposed to design an energy-saving house may go back to the formulation and reformulation to make modifications if their designs consume too much energy. They may even start to make further observations. Corresponding to these five cognitive processes, self-regulatory behaviors are recorded in the computer log files of and are distinguished as function-related or structure-related behaviors. For example, viewing the house from different angles without a specific purpose is considered to be structural observation, while specifically viewing the sun to make a connection with the function of the house is considered to be a functional observation. The details of the categorization of all behaviors will be described in the methods section. This comprehensive SRL model in the domain of engineering design connects and balances the generality and domain-specificity of SRL processes, which is of great importance in illustrating how students self-regulate their learning (Veenman, Elshout, & Meijer, 1997).

2.3. Related empirical studies in STEM fields

Although the literature of engineering education has not intensively focused on SRL, researchers have verified its importance among engineering students, and Vogt (2008) called on engineering programs to encourage faculty members to value the significance of SRL. For example, Tynjälä, Salminen, Sutela, Nuutinen, and Pitkänen (2005) found engineering students who were good self-regulated learners used deep learning strategies and had the best academic performance. Nelson, Shell, Husman, Fishman, and Soh (2015) identified several different SRL behaviors engineering students used in their foundational engineering course (Nelson et al., 2015). As well, Koh et al. (2010) suggested that engineering students needed SRL motivation to succeed in learning based on 3D simulation.

Since few empirical studies were found in the literature of engineering education, we broadly expanded our review into the STEM fields. By reviewing empirical studies in the STEM context with a specific focus on the methods these studies employed, we found the majority of relevant studies utilized a variable-oriented approach to examine the relationship between SRL processes and performance. Dent and Koenka (2016) synthesized empirical studies that have measured SRL processes and found that SRL processes correlated with learning outcomes at different levels. These correlations varied depending on the specific process, academic subject, type of SRL measure, and type of learning outcomes. Specifically, the correlation for science was stronger (r = 0.26) than for other subjects in the meta-analysis (Dent & Koenka, 2016). The positive relationship between SRL and learning outcomes has been well-established in variable-oriented studies. A few studies attempted to further distinguish learners according to their SRL profile using a person-oriented approach. For example, Ning and Downing (2015) used a person-oriented approach relying exclusively on self-reports and identified four self-regulation profiles—“competent,” “cognitively oriented,” “behaviorally oriented,” and “minimally self-regulated.” “Competent” self-regulated learners demonstrated the highest endorsement for SRL strategies, while “minimally” self-regulated learners indicated the lowest endorsement for SRL strategies. Ning and Downing (2015) also found “cognitive oriented” self-regulated learners reported high endorsement for cognitive strategies, but “behaviorally oriented” self-regulated learners had more trust in test-taking strategies. Going beyond self-reports, Bouchet, Harley, Trevors, and Azevedo (2013) used the interactions between students and a computer agent (e.g., note-taking, goal-setting actions) while learning biology on an intelligent tutoring system to classify the profile of SRL learners. The current study aims to extend person-oriented SRL research in STEM fields to an engineering design environment, investigating the following research questions:

(1) What homogenous behavioral clusters of students emerge when engaging in engineering design and what self-regulated characteristics do these clusters display?
(2) Do the behavioral profiles of self-regulated learners differ in terms of perceived engagement in metacognitive self-regulation?
(3) Do the behavioral profiles of self-regulated learners differ in terms of task performance and learning gains, and if so, how?

3. Method

In this study, we adopted a mixed research method: (1) we quantitatively analyzed the design behaviors of 108 participants and clustered them into self-regulation profiles using K-means cluster analysis; (2) we qualitatively rated students' science knowledge before and after participating in the design activities using a rubric; and (3) we conducted one-way ANOVAs to investigate the relationships between metacognitive self-regulation, performance difference, and the learning gains difference among different profiles of students. The sample, learning environment, measures, and data analysis are elaborated in the following.

3.1. Sample

The participants were 111 ninth-grade students from a suburban high school in the northeastern United States. This study was conducted with the approval of the institutional review board to protect the rights of human research subjects and all of the students who participated in this research completed consent firms. Three students' files were incomplete, leaving a sample of 108. Among them, fifty-five (50.9%) were female. The participants were enrolled in five physical science honors classes taught by a male teacher who had over 17 years' experience teaching physical sciences and 5 years' experience mentoring engineering design projects. According to the information provided by the school (as shown in Table 1), the majority of the school planned to attend 4-year colleges and were White.

3.2. The learning environment and learning task

Energy3D is a simulated environment for engineering design, and it supports designing, constructing, and analyzing green buildings that use solar energy (Xie, Schimpf, Chao, Nourian, & Massicotte, 2018). Energy3D (see Fig. 2) provides students with 3D modeling tools such as walls, windows, roofs, solar panels, and trees that enable students to quickly design and modify buildings (i.e., formulation and reformulation). It also provides simulations such as showing shadow, showing heliodon, and animating sun to help students understand the solar environment in which their houses would be built (i.e., observation). Moreover, Energy3D includes plenty of tools for quantitative analysis (e.g., compute energy and solar energy, and energy annual analysis) that help users evaluate the energy performance of their buildings and make relevant revisions to meet the net energy requirement (Xing, Wadholm, & Goggins, 2014). Finally, embedded prompts such as “describe your design ideas and explain why you think they are good ideas”

Table 1
Demographic information of students at the school.

Demographic categories	Sub-categories	Percentage
Gender	Female	50.9%
	Male	49.1%
Expected future pathways	4-year colleges	87%
	2-year college	8%
	Work or are unsure	5%
Ethnicity	White	76.7%
	Hispanic	4.6%
	African American	4.2%
	Multi-race	3.4%
	Native American	0.2%
	Native Hawaiian/Pacific Islander	0.2%

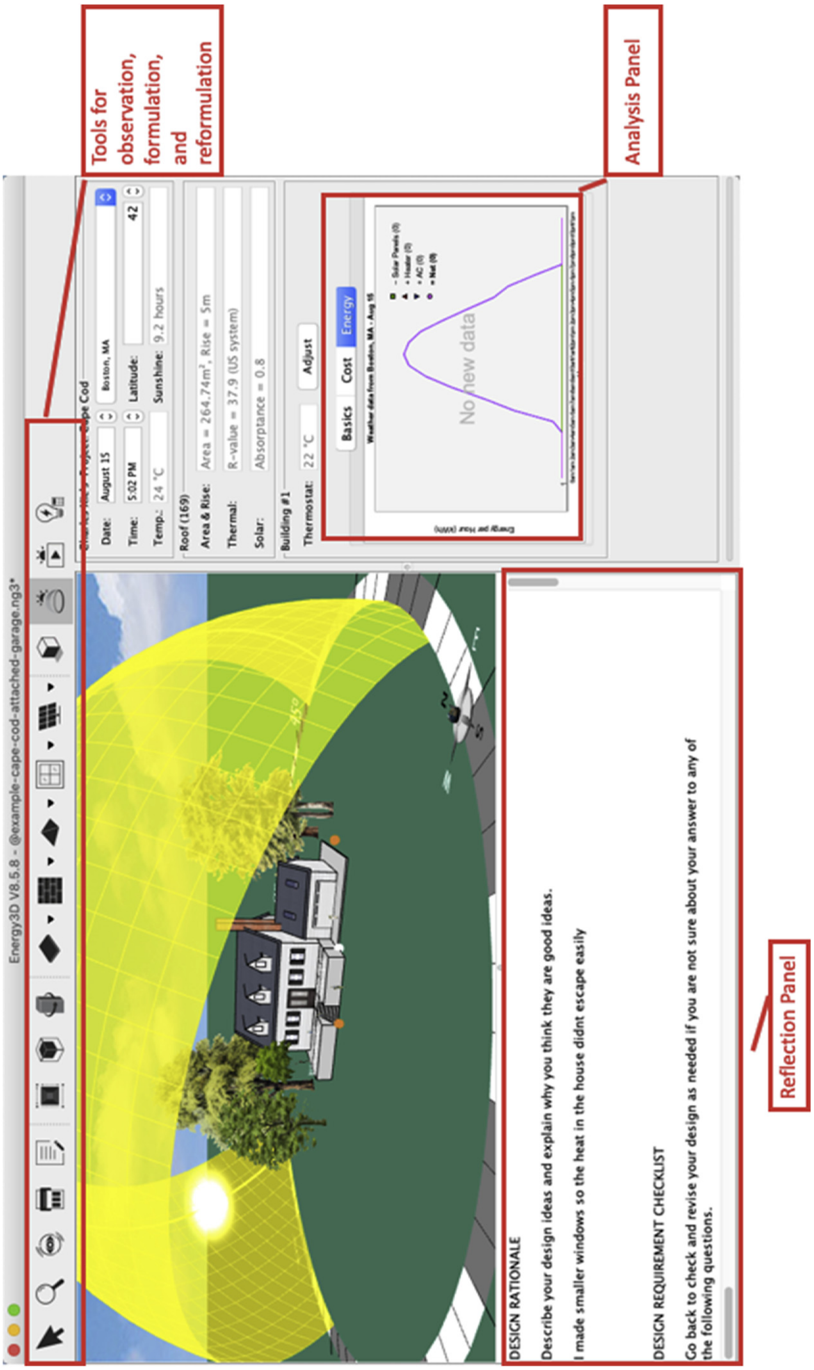

Fig. 2. Interface of Energy3D.

Fig. 3. A cape cod style house.

guided users through reflection during the learning process. All actions of users on can be recorded in Energy3D with a timeline, and the actions can indicate design processes of students and help reconstruct their self-regulation profiles.

In this study, the participants spent 50–80 min each day during a science course in which they were required to individually design an energy-plus house in Energy3D using a Cape Cod style (see Fig. 3). An energy-plus house must produce more renewable energy than it consumes over a period of 1 year. Participants were asked to complete the Cap Cord design project in two consecutive days. Furthermore, the Cape Cod house should meet the following requirements: demonstrate curb appeal; have a ratio of the total area of windows to the floor area between 0.05 and 0.15; tree trunks must be at least 2 m away from the walls of the house; roof overhang must be less than 50 cm wide. The budget for the house is $200,000. The area of the house needs to be between 100 and 150 m^2, and the height needs to be between 7 and 9 m.

Participants were given a pre-test to measure their green building science knowledge the day before the design activity. With the guidance of two researchers on the first day, students familiarized themselves with the Energy 3D platform on how to construct buildings, use embedded simulations, and perform energy analysis. Before designing the house, the students were given a two-page print instruction which specified the design requirements, listed design instructions, and important notes, and contained an engineering design cycle to guide students' design, as shown in Fig. 4. They were asked to learn the design requirements first and were encouraged to discuss their ideas with their classmates before constructing a house in Energy3D. Once the construction was done, the students could analyze the energy performance of their building using built-in analysis and revise the building to improve the energy efficiency of the houses based on the analysis results. Apart from the printed instructions, the students received minimal explicit guidance. After completing the task, the students received the same green building science knowledge test as the post-test, the metacognitive self-regulation questionnaire to measure their perceived engagement in self-regulation, and a demographic survey to identify their gender and ethnicity.

3.3. Measures

3.3.1. Learning gains of science knowledge

A green building science test was developed to assess the scientific knowledge of students in this study. This 18-question test was drawn from green building science textbooks (Hens, 2011; Montoya, 2010) and was selected based on learning opportunities provided by Energy-Plus Home design. A panel of green building science experts, high school science teachers, engineering design professors, and learning scientists reviewed the items to ensure they were appropriate and valid. The questions were related to concepts of four target domains: sun path and insolation, spatial and geometric, and heat transfer and representations. For each question, the participants were required to make a choice among design alternatives of given situations and to explain why they chose

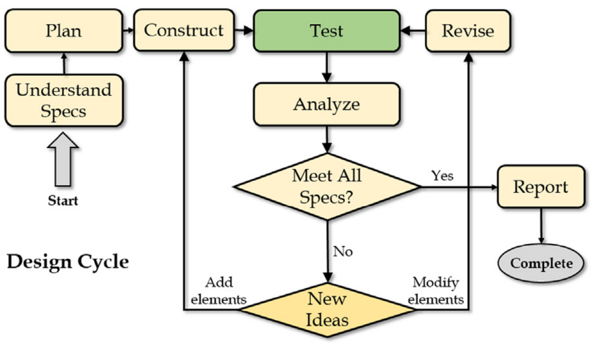

Fig. 4. The design cycle provided to the participants to guide their design.

a certain answer.

The explanations students gave for each question on the science knowledge test were scored using a 5-level scoring rubric. First, the ideas in students' responses were identified and categorized into normative, alternative, and irrelevant. Normative refers to ideas that are both scientifically correct and contribute to the ideal response. Alternative refers to ideas that are either scientifically incorrect or do not contribute to the ideal response. Irrelevant ideas include vague statements, misunderstanding of the questions, and so forth. Second, based on the numbers of normative, alternative, and irrelevant ideas, responses were categorized into five levels: Level 4 responses contain three or more connected normative ideas without any alternative ideas; Level 3 responses include two connected normative ideas and no more than one alternative idea; Level 2 responses contain one normative idea and no more than two alternative ideas; Level 1 responses only include relevant ideas and no alternative ideas; and Level 0 responses only consist of irrelevant ideas or no answers at all. For each response, the corresponding score was assigned, with a level 4 response assigned 4 points and a level 0 response zero point. Three researchers independently scored 20% of the responses to establish a substantial inter-rater reliability ranging from 0.94 to 1. The internal consistency has also been confirmed in our previous work: Cronbach's alpha was 0.82 for the pretest and 0.83 for the post-test (Chao et al., 2017). Individual learning gains of science knowledge was calculated by subtracting each participant's pretest score from the post-test score.

3.3.2. Energy performance

Students' energy performance was measured using the net annual energy of the Cape Cod they ultimately built. The net annual energy of a house equals its annual consumption energy minus its annual production energy. A negative value indicates an energy-plus house. The lower the net energy value, the more energy efficient the house.

3.3.3. Metacognitive self-regulation

The metacognitive self-regulation questionnaire was adapted from the motivated learning strategies questionnaire (MSLQ, Pintrich, Smith, Garcia, & McKeachie, 1993) to measure students' perceived engagement in self-regulation. The questionnaire consists of eight 5-point Likert items (e.g., “When designing my buildings, I make up questions to help focus on my designing”) in which 1 indicates “strongly disagree,” 3 is “neutral,” and 5 indicates “strongly agree.” The Cronbach coefficient alpha for the metacognitive self-regulation is 0.73.

3.4. Data analysis

The actions recorded in computer log files were categorized and grouped based on the domain-specific SRL model (Fig. 1). For example, students can use general view, spin view, top view, and show axes to view the structure of their design. These actions were grouped and coded as observation applying structural knowledge. Showing heliodon, showing window, animating sun, and showing annotation are the actions students may use to observe the energy-related factors, which were coded as observations applying functional knowledge.

To answer the first research question, we performed K-means cluster analysis to profile different self-regulated behavioral patterns engaging in computer-based STEM learning. The aim of cluster analysis is to identify groups of objects that have similar properties and characteristics (Hair, Black, Babin, Anderson, & Tatham, 1998). The identified cluster should exhibit high within-group homogeneity and high between-group heterogeneity (Xing, Wadholm, Petakovic, & Goggins, 2015). Based on this technique, it is possible to profile the different self-regulated behavioral patterns. An important step of clustering is to define the cluster elements. In this study, the cluster elements are the various actions in the log data (e.g., spin, showing annotation, animating sun, etc.) As a result, the data matrix has a very high dimension (108 students × 95 dimensions) for analysis. However, the K-means cluster is very sensitive to high-dimensional data, which can significantly compromise the clustering performance (McCallum, Nigam, & Ungar, 2000). Therefore, before conducting clustering analysis, principle component analysis (PCA) is performed to reduce dimensionality. PCA is a classical linear technique to reduce high dimension in the machine learning field (Hinton & Salakhutdinov, 2006). In essence, PCA performs a linear mapping of the high dimensional data to a lower dimensional space where the variance of the data in the low-dimensional representation is maximized (Wang, Xing, & Laffey, 2018).

In addition, as an unsupervised machine learning algorithm, K-means requires defining the preferred number of clusters in advance. In order to determine the optimal number of clusters used in this research, we computed the Ball statistic (Milligan & Cooper, 1985). The Ball statistic is a classic measure to compute the best number of clusters. It is used to gauge the dispersion of the data points within a cluster and between the clusters so that the data have the largest difference between clusters and smallest difference within clusters (Milligan & Cooper, 1985). Clustering analysis has the best performance when cluster K is set at the largest value of the successive difference of the Ball index values.

In sum, PCA was first used for high dimension reduction, and then the Ball statistic was calculated to determine the optimum number of clusters, and finally a K-means clustering algorithm was implemented. As cluster elements are grouped based on their similarities or the distance between them, Squared Euclidean distance (Dorling, Davies, & Pierce, 1992) was used in our study to calculate the distance between clusters.

To respond to the second research question, we conducted a one-way ANOVA to examine the relationship between various behavioral self-regulation groups and students' perceived engagement of metacognitive regulation. Similarly, regarding the third research question, we conducted a one-way ANOVA to investigate the performance difference and the learning gains difference among different profiles of students, separately. The final design products were also selected to enrich our analysis results.

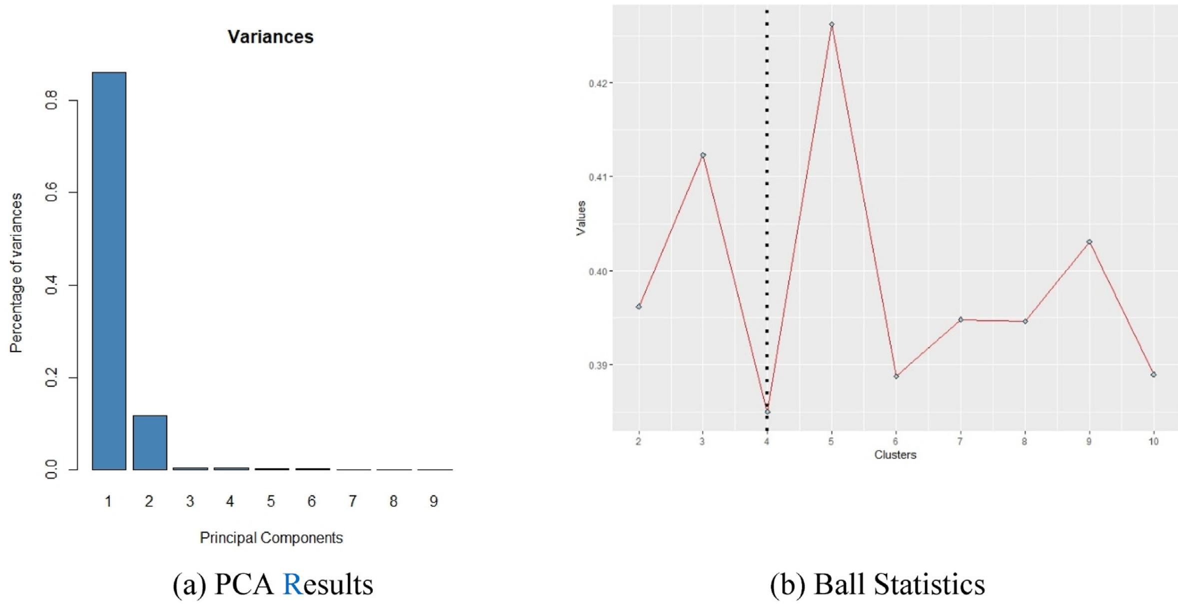

Fig. 5. PCA and Ball statistics results. (a) PCA Results. (b) Ball Statistics.

4. Results

4.1. Self-regulated behavioral profiles

The optimal number of clusters was chosen based on the following steps. Firstly, the principle component analysis was performed to reduce the high dimensionality. The result indicates that the first two components alone can explain more than 99% of the variance in the dataset as shown in Fig. 5(a). Based on these components, the Ball statistic was then calculated to identify an optimal number of clusters. The result of the Ball statistic (see Fig. 5(b)) shows that the K-means clustering has the best performance if the number of clusters is set to four. Accordingly, the result of four clusters was generated using the K-means clustering algorithm.

The final 4-cluster solution is displayed in Table 2. Following the guidelines established by Wormington and Linnenbrink-Garcia (2017) for characterizing and labeling profiles in a person-oriented approach, we identified four unique self-regulated behavioral profiles based on the most salient features and extreme low or high variables in each profile. These four profiles are: “competent self-regulated learners,” “minimally self-regulated learners,” “cognitive-oriented self-regulated learners,” and “reflective-oriented self-regulated learners.”

As shown in Table 2, the first and second profiles contain extreme high and low variables. In the first profile (n = 13, 12.15%), students demonstrated the greatest effort in the process of formulation and reformulation, especially while using functional knowledge to achieve the goal of building an energy-plus house. For example, they displayed the highest level of adding trees, adding energy, editing door, wall, window, floor, and external factors among all the students of the four clusters, indicating their competence in constructing and refining their houses. Compared with students in other profiles, they appropriately distribute their efforts across all SRL activities. They focused on constructing and modifying their house but did not spend much of their effort on the process of observation and analysis. Thus, we label this group the “competent self-regulated learners” to reflect their competency in allocating

Table 2
The four clusters categorized based on behavioral actions and their mapping with self-regulation types.

SRL	Indicators	Clusters			
		Competent (n = 13)	Minimally (n = 43)	Cognitive-oriented (n = 19)	Reflective-oriented (n = 33)
Observation	Structural (e.g., spin view, top view)	1157.46	413.98	1487.47	787.03
	Functional (e.g., show shadow, show heliodon)	20.00	5.09	24.16	9.76
Formulation	Structural (e.g., add wall, add roof)	15.31	10.47	20.32	12.45
	Functional (e.g., add energy, add trees)	127.46	53.37	98.05	76.42
Reformulation	Structural (e.g., edit wall, edit floor)	54.62	27.21	43.79	35.03
	Functional (e.g., edit energy, edit trees)	351.08	91.19	292.21	206.48
Analysis	Analysis (e.g., compute energy)	104.31	29.74	134.21	83.00
Evaluation	Evaluation (e.g., make notes)	321.69	50.74	196.58	444.00

efforts in SRL. In contrast with the first profile, the second profile (n = 42, 39.25%) had the lowest frequency on all activities, suggesting that they minimally regulated their design. For this reason, the students in the second profile were labeled "minimally self-regulated learners."

The third and fourth profiles have more moderate variables. The students in the third profile (n = 19, 17.76%) shared some commonalities with "competent self-regulated learners": they spent numerous efforts on formulation and reformulation to construct and refine their design. However, in contrast to the "competent self-regulated learners," the students in the third profile conducted the highest level of observation and analysis among the four profiles, indicating their focus on the function of their design and their efforts in understanding and analyzing their work during the process of design. They tended to maximize their cognition on the task by taking full advantage of the observation and analysis tools supported by the learning environment. This profile was considered as "cognitive-oriented self-regulated learners" to reflect their emphasis on cognition throughout the design process. The students in the fourth profile demonstrated comparatively less effort in the process of observation, formulation, reformulation, and analysis than "competent self-regulated learners" and "cognitive-oriented self-regulated learners," but more efforts than "minimally self-regulated learners." More interestingly, students in this group spent the most time reflecting on and evaluating their design. Accordingly, we label this group as "reflective-oriented self-regulated learners."

4.2. Perceived self-regulation differences among self-regulated behavioral profiles

An ANOVA was performed with cluster membership as the independent variable and perceived self-regulation level as the dependent variable. To ensure statistical power, we first conducted a power analysis using medium– to large–effect size 0.35. For a four-group one-way ANOVA, only 96 subjects are required. Therefore, in our study, 108 students should provide a large enough sample size to generate reliable findings. The results of ANOVA show a significant difference with about medium effect size according to Miles and Shevlin (2001) among these four self-regulated behavioral groups in terms of perceived self-regulation ($F = 4.10$, $p = .046 < 0.05$, $\eta^2 = 0.05$). As shown in Table 3, the competent self-regulated learners showed the highest level of perceived self-regulation ($M = 38.23$), which is not surprising. The competent self-regulated group not only exhibited the highest level of self-regulated actions, but they also were aware of their high self-regulation behavior. Surprisingly, the minimally self-regulated group perceived their self-regulation ($M = 35.90$), as second only to the competent self-regulated group. That is, it was higher than that of the cognitive-oriented self-regulated ($M = 32.95$) and the reflective-oriented regulated groups ($M = 33.79$). Among the four groups, the cognitive-oriented self-regulated group perceived they had the lowest level of meta-cognitive self-regulation.

4.3. Task performance and learning gains differences among self-regulated behavioral profiles

A one-way ANOVA was performed with cluster membership as the independent variable and net energy performance as the dependent variable. Results indicated that there is a significant difference with medium effect size in the net annual energy of the houses built by students of the four self-regulated groups ($F = 6.40$, $p = .01 < .05$, $\eta 2 = 0.06$). As shown in Table 3, the reflective-oriented self-regulated group performed best in terms of energy efficiency. As aforementioned, this group of students spent most efforts on reflection and evaluation processes by adopting embedded prompts to guide design. Therefore, it is not surprising that they performed best in terms of net energy of the houses they built. Similarly, the average net energy of the houses built by the cognitive-oriented self-regulated group was negative, which met the energy requirement of the design task. As described above, the students in the cognitive-oriented self-regulated group conducted the energy analysis most frequently when designing. Moreover, the minimally self-regulated learners, not surprisingly, performed worst. Their average energy value was farthest away from the net energy requirement among the four groups of students. However, contrary to our expectations, the competent self-regulated learners did not meet the energy-plus goal.

Similarly, a one-way ANOVA was performed with science knowledge learning gains as the dependent variable. No significant difference was found among the four groups. However, the descriptive statistics showed that the competent self-regulated learners had the greatest learning gains and the minimally self-regulated learners had the least learning gains.

Examples were further drawn from the four profiles to illustrate their differences in terms of final products. The final products not only reveal students' application of functional knowledge as indicated in net energy value, but also reflect their utilization of structural knowledge. As displayed in Fig. 6, one of the minimally self-regulated learners (e.g., E14) obviously was not engaged in SRL, and as a result basic structures (e.g., windows, doors) were missing in his design and net energy value was positive (Net value = 46200). In contrast, the reflective learner (e.g., F02) had clearly focused on the energy-plus function in his design. This is

Table 3
Differences among four clusters on perceived metacognitive self-regulation and net energy performance on science knowledge learning gains.

Dependent variable	Competent	Cluster Mean			
		Minimally	Cognitive-oriented	Reflective-oriented	ANOVA F value
Perceived metacognitive self-regulation	38.23	35.9	32.95	33.79	4.10*
Net energy performance	1361.35	2955.2	−575.73	−1732.84	6.40*
Learning gains of science knowledge	20.46	16.52	17.63	18.21	0.002

*p < .05.

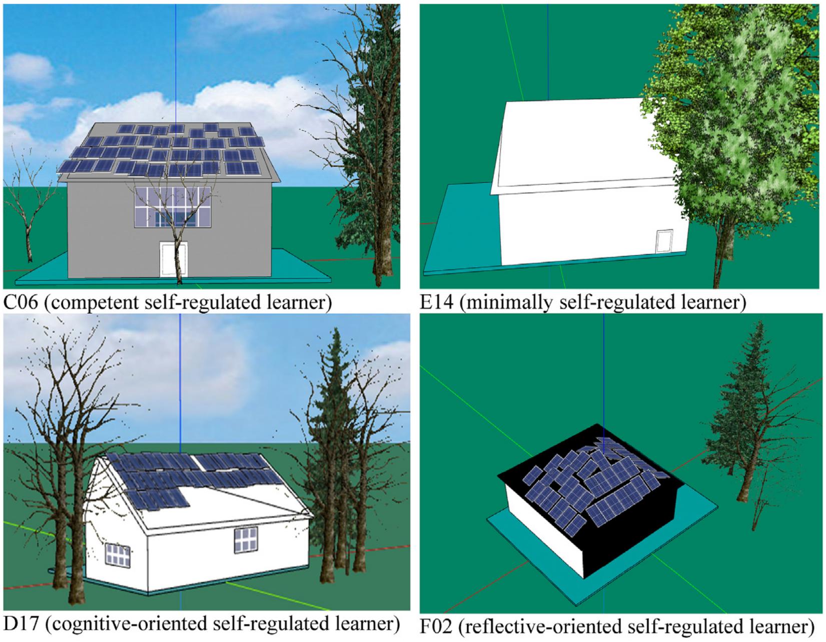

Fig. 6. Sample product of students in four profiles.

why he put a lot of solar panels on the roof even though half of them were not producing energy. As for the competent (e.g., C06) and cognitive-oriented self-regulated learner (e.g., D17), they both paid attention to the structure of the house and intentionally arranged the solar panel towards one direction. But the competent learner was more determined in his design, while the cognitive-oriented learner relied more on the observation and analysis supported by the learning environment. The competent learner added trees to block every window, but the cognitive-oriented learner only put trees based on her observation of the sun. Thus, it is not surprising that the competent self-regulated learner did not achieve a good net energy value (Net Value = 161). To sum up, these four examples partly support the aforementioned findings on the task performance and learning gains, which will be further discussed in the following section.

5. Discussion

This study contributes to the advancement of SRL theories in STEM education. We identified four behavioral self-regulation profiles: competent, minimal, cognitive-oriented, and reflective self-regulated learners. On one hand, these findings further confirm the existence of two extreme self-regulated learners identified by prior researchers (e.g., Barnard-Brak, Lan, & Paton, 2010; Ning & Downing, 2015): competent self-regulated learners who represent the highest SRL and minimally self-regulated learners who represent the lowest SRL. On the other hand, these four profiles provide a complement to Ning's et al. (2015) research findings, wherein a profile is characterized by self-reported SRL strategies. Specifically, the identification of competent, cognitive-oriented, and reflective-oriented learners further extends SRL theory in the STEM field, providing empirical support for the assessment of SRL in engineering design. Self-regulated learners exhibit distinct SRL behaviors (Barnard-Brak et al., 2010), and the current study shows that these behaviors are closely related to SRL processes. SRL processes in engineering design (i.e., observation, formulation, reformulation, analysis, and evaluation) emerged as factors that quantitatively differentiate SRL profiles. Moreover, cognitive-oriented learners were more engaged in observation and analysis (i.e., application of functional knowledge) than competent self-regulated learners. This small difference indicates that some learners tend to devote themselves to the function of their design (Carlsen, 1998), which implies educators should assess both the functional and structural knowledge of students. More notably, the identification of a reflective-oriented self-regulated learner adds to existing evidence that experienced designers constantly engage in reflection to evaluate the design product (Adams, Turns, & Atman, 2003). As discussed above, students display different self-regulated behavioral profiles in engineering design, and the characteristics of each profile can be more salient when computer log files are used in a person-oriented approach.

With regard to the perceived self-regulation differences among the four identified profiles, we found competent self-regulated learners had a good self-awareness of their SRL and perceived themselves as the best SRL learners. However, cognitive-oriented self-regulated learners underestimate themselves, and minimally self-regulated learners overestimate themselves. These findings suggest that students have different levels of self-awareness of their SRL (Zimmerman, 2002). Self-awareness is inversely associated with self-regulation failure. As such, strategies such as asking students to self-record their learning process and then review the process can be taken into account to increase students' awareness of their SRL (Hadwin & Oshige, 2011). In addition, visualizations or a dashboard of students' behavioral profiles established in computer-based learning environments may also help students improve their self-awareness.

Results from our third research question reveal that task performance is not always consistent with learning gains in terms of the differences among four profiles. Competent self-regulated learners had the greatest learning gains, but they did not perform best in the task. Reflective-oriented learners surprisingly performed best in the task. A further examination of the final design product of one reflective-oriented self-regulated learner suggests that he is more concerned about the functional aspect of his design. It is possible that students in this group constantly reflect and evaluate their functional design to achieve best results in the task. However, competent self-regulated learners are more engaged in regulating their learning and acquiring knowledge, which may at times harm their task performance. This is why students in this profile have the most learning gains. Similar to prior studies that examine the correlation between SRL and learning outcomes (Dent & Koenka, 2016), this study provides evidence that the relationship between SRL and learning outcomes varies depending on the task and the assessment of learning outcomes. Therefore, a more comprehensive assessment of student learning may yield changes in students' SRL processes.

6. Conclusions and implications for STEM education

This study highlights the importance of analyzing SRL processes within the STEM field and sets a good example for employing a person-oriented approach on behaviors recorded in computer log files. Four distinct SRL behavioral profiles are identified and examined and they are related to students' perceived self-regulation, task performance, and learning gains. The findings in this study lend empirical to support the self-regulation framework, and they reveal the importance of self-regulation to student performance in STEM learning. In addition, the computer trace data that was used in the learning analysis could be used to provide learners with individualized prompts or feedback in response to their self-regulation profiles. However, this study is also limited in generalizability as it is based on the learning environment (Energy 3D) specifically developed for engineering design.

The findings of this study have various implications for STEM teaching and learning. First of all, teachers and computer programs need to provide individualized scaffolding and instructions to students with various SRL characteristics. For example, prompts or strategies can be developed to motivate students who spend the least effort and minimally regulate themselves. Extra interventions, such as helping them set up personal goals, would enhance their engagement in learning integrated STEM concepts and master STEM skills accordingly. Effort can be made to guide the cognitive-oriented self-regulated learners towards competent self-regulated learners through highlighting the importance of comprehensive knowledge in STEM projects. Furthermore, we strongly recommend that educators design learning environments that can improve students' self-awareness of SRL since students sometimes underestimate or overestimate themselves. As suggested by a previous researcher, students may go through a calibration process before having an accurate estimation of their SRL (Stone, 2000). Thus, dynamic visualizations or learning analytics dashboards may help students with their calibration by increasing their self-awareness of their learning processes. Finally, it is of great importance to leverage different STEM concepts in the evaluation of STEM learning, considering students performed differently in different STEM components. For example, we find that students who had the largest learning gains of science knowledge (the science component) did not perform best in saving energy (the engineering component) in their house design. Therefore, we should evaluate students from different perspectives to ensure all the STEM components are taken into consideration. This is especially true in STEM projects where cross-disciplinary knowledge is needed to teach and learn integrated STEM concepts.

7. Limitations and future directions

This study has several limitations. First, even though we have a large enough sample size for the research based on the power analysis, most subjects were white grade 9 students and all attended a single school in the northeastern United States, which limits the generalizability of the results. Second, for K-means clustering analysis, the initial seeds selection has a strong impact on the final clustering results. The scaling of the data sets can also influence the clustering results. Therefore, future research might try various initial seeds and scaling methods with different data sets to ensure the consistency of the findings.

Future research can build on this study by taking cultural backgrounds, gender, ages, socio-economic backgrounds, and other demographic factors into consideration. More participants can be recruited and our findings could be tested using tasks designed to teach other STEM fields. In addition, instead of conducting post-analysis of students' self-regulation, it would be interesting to embed this analysis in the learning environment to provide real-time feedback to students to examine how making them aware of their behavior may influence their self-regulation. Finally, the relationship between individual design performance and students' science knowledge assessment performance and how these two competencies can strengthen one another can be better understood to enhance student performance in both design and cognitive understanding.

Conflicts of interest

The authors declare that they have no conflict of interests.

Acknowledgements

This work is supported by the National Science Foundation (NSF) of the United States under grant numbers 1512868, 1348530 and 1503196. Any opinions, findings, and conclusions or recommendations expressed in this paper, however, are those of the authors and do not necessarily reflect the views of the NSF. The authors are indebted to Joyce Massicotte, Elena Sereiviene, Jie Chao, Xudong

Huang, and Corey Schimpf for assistance and suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2019.103669.

References

Adams, R. S., Turns, J., & Atman, C. J. (2003). Educating effective engineering designers: The role of reflective practice. *Design Studies, 24*(3), 275–294.
Alexander, P. A., Dinsmore, D. L., Parkinson, M. M., & Winters, F. I. (2011). *Self-regulated learning in academic domains Handbook of self-regulation of learning and performance.* New York, NY, US: Routledge/Taylor & Francis Group393–407.
Barnard-Brak, L., Lan, W. Y., & Paton, V. O. (2010). Profiles in self-regulated learning in the online learning environment. *International Review of Research in Open and Distance Learning, 11*(1), 61–79.
Bouchet, F., Harley, J. M., Trevors, G. J., & Azevedo, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *JEDM| Journal of Educational Data Mining, 5*(1), 104–146.
Buric, I., & Soric, I. (2012). The role of test hope and hopelessness in self-regulated learning: Relations between volitional strategies, cognitive appraisals and academic achievement. *Learning and Individual Differences, 22*(4), 523–529.
Carlsen, W. S. (1998). Engineering design in the classroom: Is it good science education or is it revolting? *Research in Science Education, 28*(1), 51–63.
Chao, J., Xie, C., Nourian, S., Chen, G. H., Bailey, S., Goldstein, M. H., ... Tutwiler, M. S. (2017). Bridging the design-science gap with tools: Science learning and design behaviors in a simulated environment for engineering design. *Journal of Research in Science Teaching, 54*(8), 1049–1096.
Chiu, J. L., Malcolm, P. T., Hecht, D., DeJaegher, C. J., Pan, E. A., Bradley, M., et al. (2013). WISEngineering: Supporting precollege engineering design and mathematical understanding. *Computers & Education, 67*, 142–155.
Cleary, T. J., & Callan, G. L. (2018). Assessing self-regulated learning using microanalytic methods. In D. H. Schunk, & J. A. Greene (Eds.). *Handbook of self-regulation of learning and performance* (pp. 338–351). (2nd ed.). Routledge.
Crismond, D. P., & Adams, R. S. (2012). The informed design teaching and learning matrix. *Journal of Engineering Education, 101*(4), 738–797.
Dabbagh, N., & Kitsantas, A. (2013). Using learning management systems as metacognitive tools to support self-regulation in higher education contexts. In R. Azevedo, & V. Aleven (Eds.). *International handbook of metacognition and learning technologies* (pp. 197–211). New York: Springer New York, NY.
Dasgupta, C., Magana, A. J., & Vieira, C. (2019). Investigating the affordances of a CAD enabled learning environment for promoting integrated STEM learning. *Computers & Education, 129*, 122–142.
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.
Dixon, R. A. (2010). *Experts and novices: Differences in their use of mental representation and metacognition in engineering design.* 3455660 Ph.D. Ann Arbor: University of Illinois at Urbana-Champaign. Retrieved from https://proxy.library.mcgill.ca/login?url=https://search.proquest.com/docview/868724614?accountid=12339 ProQuest Dissertations & Theses Global database.
Dorling, S. R., Davies, T. D., & Pierce, C. E. (1992). Cluster analysis: A technique for estimating the synoptic meteorological controls on air and precipitation chemistry—results from Eskdalemuir, South Scot- land. *Atmospheric Environment Part A. General Topics, 26*, 2583–2602.
Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education, 94*(1), 103–120.
Fan, S. C., & Yu, K. C. (2017). How an integrative STlEM curriculum can benefit students in engineering design practices. *International Journal of Technology and Design Education, 27*(1), 107–129.
Gero, J. S. (1990). Design prototypes - a knowledge representation schema for design. *AI Magazine, 11*(4), 26–36.
Gero, J. S., & Kannengiesser, U. (2004). The situated function–behaviour–structure framework. *Design Studies, 25*(4), 373–391.
Hadwin, A., & Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record, 113*(2), 240–264.
Hair, J. F., Jr., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis with readings.* Englewood Cliffs, NJ: Prentice-Hall.
Hens, H. (2011). *Applied building physics.* Berlin, Germany: Ernst&Sohn.
Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*(5786), 504–507.
Howard, T. J., Culley, S. J., & Dekoninck, E. (2008). Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Design Studies, 29*(2), 160–180.
Koh, C., Tan, H. S., Tan, K. C., Fang, L., Fong, F. M., Kan, D., ... Wee, M. L. (2010). Investigating the effect of 3D simulation based learning on the motivation and performance of engineering students. *Journal of Engineering Education, 99*(3), 237–251.
Lajoie, S. P., Poitras, E. G., Doleck, T., & Jarrell, A. (2015). Modeling metacognitive activities in medical problem- solving with BioWorld. In A. Peña-Ayala (Ed.). *Metacognition: Fundaments, applications, and trends*. Switzerland: Springer International Publishing.
Lawanto, O., & Johnson, S. D. (2012). Metacognition in an engineering design project. *International Journal of Engineering Education, 28*(1), 92–102.
Lewis, T. (2006). Design and inquiry: Bases for an accommodation between science and technology education in the curriculum? *Journal of Research in Science Teaching, 43*(3), 255–281.
Linnenbrink-Garcia, L., & Wormington, S. V. (2017). Key challenges and potential solutions for studying the complexity of motivation in schooling: An integrative, dynamic person-oriented perspective. *British Journal of Educational Psychology Monograph Series, 12*, 89–108.
McCallum, A., Nigam, K., & Ungar, L. H. (2000, August). Efficient clustering of high-dimensional data sets with application to reference matching. *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 169–178). ACM.
Miles, J., & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers.* Sage.
Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika, 50*, 159–179.
Montoya, M. (2010). Upper Saddle River, N.J.: Prentice Hall.
Nelson, K. G., Shell, D. F., Husman, J., Fishman, E. J., & Soh, L. K. (2015). Motivational and self-regulated learning profiles of students taking a foundational engineering course. *Journal of Engineering Education, 104*(1), 74–100.
Ning, H. K., & Downing, K. (2015). A latent profile analysis of university students' self-regulated learning strategies. *Studies in Higher Education, 40*(7), 1328–1346.
Ozcan, Z. C. (2016). The relationship between mathematical problem-solving skills and self-regulated learning through homework behaviours, motivation, and metacognition. *International Journal of Mathematical Education in Science & Technology, 47*(3), 408–420.
Peng, Y., Hong, E., & Mason, E. (2014). Motivational and cognitive test-taking strategies and their influence on test performance in mathematics. *Educational Research and Evaluation, 20*(5), 366–385.
Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement, 53*(3), 801–813.
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.
Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*(3), 421–442.
Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review, 12*(4), 437–475.

Strobel, J., Wang, J., Weber, N. R., & Dyehouse, M. (2013). The role of authenticity in design-based learning environments: The case of engineering education. *Computers & Education, 64*, 143–152.
Tynjälä, P., Salminen, R. T., Sutela, T., Nuutinen, A., & Pitkänen, S. (2005). Factors related to study success in engineering education. *European Journal of Engineering Education, 30*(2), 221–231.
Veenman, M. V. J., Elshout, J. J., & Meijer, J. (1997). The generality vs domain-specificity of metacognitive skills in novice learning across domains. *Learning and Instruction, 7*(2), 187–209.
Vogt, C. M. (2008). Faculty as a critical juncture in student retention and performance in engineering programs. *Journal of Engineering Education, 97*(1), 27–36.
Wang, X., Xing, W., & Laffey, J. M. (2018). Autistic youth in 3D game-based collaborative virtual learning: Associating avatar interaction patterns with embodied social presence. *Br. J. Edu. Technol. 49*(4), 742–760.
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.
Xie, C., Zhang, Z. H., Nourian, S., Pallant, A., & Hazzard, E. (2014). Time Series Analysis Method for Assessing Engineering Design Processes Using a CAD Tool. *International Journal of Engineering Education, 30*(1), 218–230.
Xing, W., Guo, R., Petakovic, E., & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable genetic programming: integrating learning analytics, educational data mining and theory. *Computers in Human Behavior, 47*, 168–181.
Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17.
Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. *Handbook of Self-Regulation,* 13–39.
Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64–70.
Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183.
Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135–147.

Juan Zheng is a PhD candidate and research assistant at Educational Counselling and Psychology (ECP), McGill University with background in educational technology and learning sciences. Her research interests are self-regulated learning, academic emotions, and educational data mining

Wanli Xing is an Assistant Professor of Educational Technology at University of Florida. His research interests are artificial intelligence, learning analytics, STEM education and online learning.

Gaoxia Zhu is a PhD candidate and research assistant in the Institute for Knowledge Innovation & Technology, Ontario Institute for Studies in Education (OISE), University of Toronto with background in Educational Technology and Curriculum Studies. Her research interests include Knowledge Building, learning analytics, and CSCL.

Guanhua Chen is a postdoctoral researcher at the Concord Consortium. His research interests are science and engineering education research, computational thinking, machine learning, and educational data mining.

Henglv Zhao is a PhD student in Educational Psychology at Texas Tech University. He earned his bachelor and master's degree in material science and engineering. His research interests include educational data mining, artificial intelligence and machine learning.

Charles Xie is a senior scientist at the Concord Consortium. His research includes computational science, learning science, machine learning, artificial intelligence, computer-aided design, mixed reality, Internet of Things, molecular simulation, solar energy engineering, and infrared imaging. He develops the Energy3D software that supports this research.