JOURNAL OF GEOSCIENCE EDUCATION
https://doi.org/10.1080/10899995.2024.2384340



# Using artificial intelligence teaching assistants to guide students in solar energy engineering design

Q5 Shannon Sung, Xiaotong Ding, Rundong Jiang, Elena Sereiviene, Dylan Bulseco and Charles Xie Q1

Institute for Future Intelligence, Natick, Massachusetts, USA

**ABSTRACT**

Engineering projects, such as designing a solar farm that converts solar radiation shined on the Earth into electricity, engage students in addressing real-world challenges by learning and applying geoscience knowledge. To improve their designs, students benefit from frequent and informative feedback as they iterate. However, teacher attention may be limited or inadequate, both during COVID-19 and beyond. We present Aladdin, a web-based computer-aided design (CAD) platform for engineering design with a built-in artificial intelligence teaching assistant (AITA). We also present two curriculum units (Solar Energy Science and Solar Farm Design), where students explore the Sun-Earth relationship and optimize the energy output and yearly profit of a solar farm with the help of the AITA. We tested the software and curriculum units with over 100 students in two Midwestern high schools. Pre- and post-survey data showed improvements in understanding of science concepts and self-efficacy in engineering design. Pre-post analysis of design performance gains reveals that AI helped lower achievers more than higher achievers. Interviews revealed students' values and preferences when receiving feedback. Our findings suggest that AITAs may be helpful as an additional feedback mechanism for geoscience and engineering education. Future efforts should focus on improving the usability of the software and providing multiple types of feedback to promote inclusive and equitable use of AI in education.

**ARTICLE HISTORY**
Received 9 August 2022
Revised 9 July 2024
Accepted 19 July 2024

**KEYWORDS**
Engineering design; artificial intelligence; renewable energy; feedback; technology enhanced learning Q3

## Introduction

There is an increasing demand to integrate engineering design with geoscience education. At the K-12 level, the Next Generation Science Standards (NGSS) listed seven Earth and Space Science (ESS) performance expectations that incorporate engineering practices. For example, high school students are expected to "evaluate competing design solutions for developing, managing, and utilizing energy and mineral resources based on cost-benefit ratios" (NGSS Lead States, 2013). In addition, the K-12 science faculty, including ESS educators, share the responsibility to address 14 separate NGSS performance expectations for engineering design, such as "*design[ing] a solution to a complex real-world problem by breaking it down into smaller, more manageable problems that can be solved through engineering*" (NGSS Lead States, 2013). According to the National Research Council's *A Framework for K-12 Science Education*, on which the NGSS is based, a major advantage of integrated science and engineering education is that "*[f]rom a teaching and learning point of view, it is the iterative cycle of design that offers the greatest potential for applying science knowledge in the classroom and engaging in engineering practices*" (NRC, 2012, pp. 201–202). The interactivity of and repeated involvement in engineering design projects may also help trigger and maintain students' situational interest in ESS (van der Hoeven Kraft, 2017).

Of all engineering design projects within an ESS context, renewable energy engineering may be one of the most familiar to a K-12 audience. Take solar energy engineering—the design and deployment of solar power systems—for an example. Prior research suggests that while an overwhelming majority of students reported some familiarity with the concept of solar panels and many reported seeing them in their everyday lives, much fewer could use ESS knowledge such as solar angles to explain what time of day solar panels worked best (Kishore & Kisiel, 2013). Therefore, a solar energy design project can both relate to students' personal experiences with solar energy and reinforce their ESS knowledge through repeated application in an iterative design process. For example, how can the design of utility-scale solar panel arrays—or solar farms—integrate ESS with engineering? For starters, the energy output of solar panels depends on solar irradiance, which fluctuates according to the Sun's position in the sky. An optimal solar farm design will use an appropriate tilt angle to maximize the solar insolation (the total incident solar radiation across a certain time) and thus the energy output. In addition, the exact energy output of a solar farm is dependent on a number of other factors, such

**CONTACT** Shannon Sung shannon@intofuture.org 26 Rockland St., Natick, MA 01760, USA.
Curriculum & Instruction

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10899995.2024.2384340. Q2

as the latitude (which determines the daytime length and the Sun's relative position), the local weather (which determines the number of sunshine hours and local temperature), and air pollution (which can absorb and scatter light), all of which must be accounted for in an accurate yield analysis.

Solar energy engineering education can be a powerful response to the critical issues that the global coronavirus pandemic brought to light as educators in all disciplines were forced to shift toward virtual learning tools and online teaching. Geoscience educators were acutely aware of the need to reduce barriers for disadvantaged students but saw those new factors, such as internet access and family dynamics impacted education (Riggs, 2020). Other critical issues also surfaced due to the pandemic such as the climate crisis and how both tragedies disproportionately impact marginalized communities (Behune, 2020). The Biden Administration signed an executive order to decarbonize the energy sector (The United States Government, 2021), and the amount of renewable energy generated reached a record high of 28% in April 2022 (U.S. Energy Information Administration, 2022), which demonstrates the importance of solar energy engineering and education in this sector. Meeting the decarbonization goal requires a solar workforce of as many as 500,000–1,500,000 people by 2035 (U.S. Department of Energy, 2021), which serves as a reminder that engineering design projects should be integrated into regular ESS education so that students can be prepared to apply their geoscience knowledge to mitigating global challenges such as the climate crisis.

The engineering design process is iterative, and improvement is incremental, meaning that students would typically require frequent feedback on their design process and product, often from either their teachers or their peers, so that they can evaluate the pros and cons of their current design, assess their application of scientific principles, and explore potential next steps. Unfortunately, teachers are often unable to look over each student's shoulder to provide individual feedback on each design iteration due to a lack of time or expertise (An & Mindrila, 2020). Peer feedback may be more available but not necessarily as effective without proper training. The situation was exacerbated by the total interruption of all face-to-face interactions at the height of the pandemic, meaning that students were often left with no feedback during their learning.

In addition to introducing new challenges, the COVID-19 crisis also highlighted existing shortcomings in science and engineering education, especially around issues of equity and inclusion. For example, the cost of physical materials can be a barrier for students with low socioeconomic status, limiting their access to and success in engineering design projects. Traditional engineering projects may not be accessible for students with chronic illness or disabilities, who may rely more on virtual learning than their peers (Porter et al., 2021; Thornton et al., 2022). Also, some students may not actively seek teacher or peer feedback due to their personality or neurodiversity. In each case, the lack of alternatives may discourage certain students from developing an interest or expertise in science and engineering. Therefore, the necessity of an ever-available virtual option has become evident to students and teachers. Alternative feedback mechanisms need not replace all in-person teacher and peer feedback. Still, they can serve as a safety net and allow students to personalize their learning based on their diverse needs.

Recent developments in artificial intelligence (AI) have propelled a wave of educational applications in assessment, tutoring, and feedback (Afzaal et al., 2021; Darvishi et al., 2022; Goldin et al., 2017; Hooda et al., 2022; Mirchi et al., 2020; Porter & Grippa, 2020). In the field of engineering design, AI has been used in computer-aided design (CAD) and computer-aided engineering (CAE) settings (Shu et al., 2019; Yoo et al., 2021), computational geoscience (Bergey, 2020), and renewable energy engineering (Vahdatikhaki et al., 2022). There has been some exploration of its capability to assess engineering design performance (Xing et al., 2021), but little has been reported about its potential as a feedback mechanism in engineering education.

To advance inclusive and equitable science and engineering education and promote student agency in developing solutions to global challenges using geoscience knowledge, we introduce 1) a virtual platform for engineering design called Aladdin (Figure 1); 2) a built-in artificial intelligence teaching assistant (AITA) capable of providing individual design feedback, and; 3) a week-long solar energy science and engineering curriculum. Aladdin is an integrated CAD and CAE tool for renewable energy engineering (Xie et al., 2023). The design of the AITA was informed by the field of heuristics (Gigerenzer, 2008), which has a long tradition in math teaching (Higgins, 1971; Hughes, 1974; Lucas, 1974), has been observed as a scaffolding technique for teaching assistants (Radford et al., 2014), and was viewed as a suitable solution for AI agents (Al-Shaery et al., 2022).

The week-long curriculum consists of two units: In the first unit, "Solar Energy Science," students explore basic ESS concepts related to solar energy engineering, such as solar angles and the projection effect, which describes the varying angles of the Sun that shine on a designated surface will affect the amount of solar energy it gets. In the second unit, "Solar Farm Design," students explored the design requirements of a solar farm, followed the engineering design process to create their own solar farm designs, and used the AITA to improve their designs. We also present our evaluation of the software and curriculum using data from a recent field study and discuss limitations and future opportunities for AITAs in geoscience education.

## Study population and setting

The study took place in May 2022 in two suburban high schools in a Midwestern state. The demographics are reported in Table 1. Students from School 1 and School 2, which have similar demographics, took environmental science and physical science, respectively. Both schools had resumed in-person learning at the time of the study. In School 1, students sat individually and used their own laptops. In School 2, students sat in groups of one to four people, each assigned to use a school-issued Chromebook. Even though the participants were encouraged to interact with

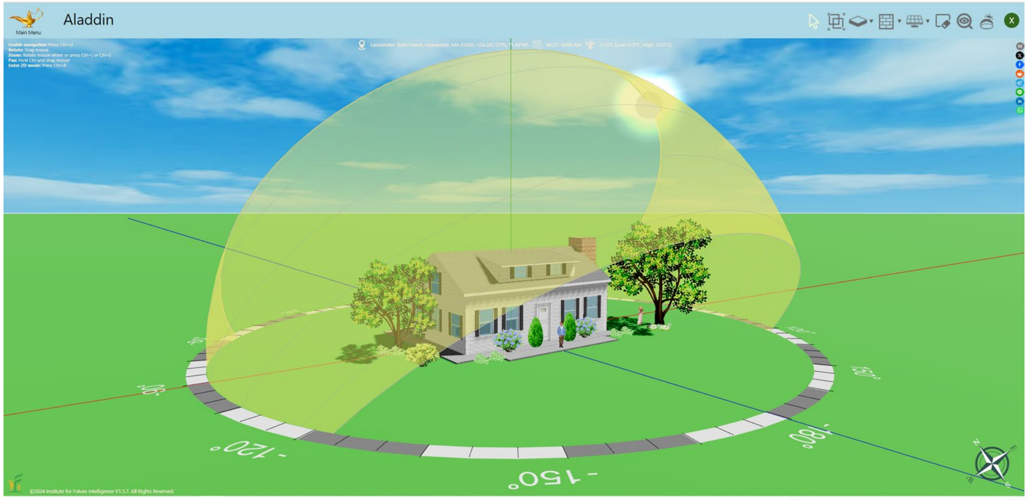


**Figure 1.** A screenshot of the virtual heliodon feature in the Aladdin software. The virtual heliodon visualizes the Sun's current position relative to an observer on Earth and the Sun's possible paths throughout a year for a given location. Students can input any date, time, and latitude into Aladdin, and the heliodon visualization will update automatically to reflect the change. Students can also turn on the animation feature to see the Sun move across the sky in a day.

**Table 1.** The demographic information of both schools that participated in the study.

| | School 1 | School 2 |
|---|---|---|
| Number of periods | 4 | 2 |
| Total number of students in each period | 25, 32, 28, 26 | 16, 15 |
| Total number of students enrolled | 111 | 31 |
| Subject | AP Environmental Science | Physical Science |
| Age | 15–18 (mode: 17) | 16–18 (mode: 17) |
| Gender | Female: 45.0%<br>Male: 39.6%<br>Didn't report: 12.6%<br>Prefer not to answer: 2.7% | Female: 38.7%<br>Male: 54.8%<br>Prefer not to answer: 6.5% |
| Ethnicity (only showing those with population > 5%) | White/Caucasian: 55.0%<br>Asian/Pacific Islander: 20.7%<br>Didn't report: 12.6% | White/Caucasian: 61.3%<br>Multiple ethnicity: 16.1%<br>Hispanic American: 6.5%<br>Black or African American: 6.5% |

their teacher, peers, and AITA, each student was expected to fill out the AI worksheet and complete the AITA curriculum individually. Two science teachers—one from each school—participated in three hours of professional development before implementing the Solar Farm Design curriculum in their classrooms. The resources covered during the professional development can be found in the following "Materials and Implementation" section.

## Materials and implementation

Before the implementation, the teachers received access to the free Aladdin software (http://intofuture.org/aladdin.html), the Solar Energy Science unit (http://intofuture.org/aladdin-solar-science.html), and the Solar Farm Design unit (http://intofuture.org/aladdin-solar-farm-design-ai.html). The units included student worksheets, teacher guides, design journals, and links to pre-made Aladdin models. Students and teachers could run Aladdin directly in the browser using their Chromebooks or laptops. All worksheets and surveys were also completed online using Google Suite. The teachers had editor access to all Google Docs files and could view and leave comments on student worksheets. They could also view the Google Form responses of the pre- and post-surveys.

The full curriculum took five to seven days to implement. On the first day, a pre-survey was administered in class, which took about 10 min. The teacher then introduced the first unit, "Solar Energy Science," which had also been used in other curriculum projects (Sung et al., 2022). Over the next 2–3 days, students worked through the solar energy science unit in a self-directed fashion. The main learning objectives of this section were to describe the Sun's position using solar elevation and azimuth angles, describe the daily and seasonal changes of solar angles, describe the relationship between the angle of incidence and the energy output of a solar panel, and explain the optimal tilt angles of a solar panel that maximizes the energy output in each season and in a year. Each activity followed the Predict-Observe-Explain framework (White & Gunstone, 2014). For example, students would first predict the best tilt angle for fall, then conduct an investigation in Aladdin, where they compared the simulated daily energy output of solar panels with different tilt angles. Finally, they were asked to explain this result using the solar energy science concepts they learned earlier, such as the solar elevation angle and projection effect. At the end of the unit, students completed a challenge called "Optimize It!" where they needed to find the best position and angle to place a single solar panel in a yard surrounded by trees, such that the panel would generate the most yearly output.

After students completed the Solar Energy Science unit, they continued to the Solar Farm Design unit, where they were tasked with designing a solar farm that would generate the most yearly profit for their town. Here, the main learning objectives were to evaluate a design solution using the given criteria and constraints, collect evidence of the design performance using computer simulations, improve the design performance through iterations, create a design that meets the given criteria and constraints, and explain the choice of design variables using scientific principles. Students were first introduced to the design criteria, constraints, and variables of a solar farm. The design process began with the students evaluating an existing solar farm design with sub-optimal performance: a negative yearly profit. Students brainstormed how they could change the three design variables—tilt angle, row width (RW), and inter-row spacing (IRS)—to improve the performance and were asked to document their reasoning. After choosing one design variable to change and specifying its new value, students input the new design variable into Aladdin's layout wizard, which automatically updated the solar farm design layout based on the specified variables. Students were directed to save their new design as a new file on the cloud storage as a method of showcasing their design artifacts and version control. Students then used Aladdin to calculate the yearly energy output and profit of their new design and compare it with the performance of the previous iteration, and they reflected on their learning during this iteration. An example iteration was provided on the design instruction to help explain the design process and clarify the expectation of student responses. The teacher also demonstrated how to go through a design iteration in Aladdin on the projector screen. After that, students were given at least one class period to create their own solar farm designs, and they were directed to document their full design process, including their design variables, performances, and reflections, in a pre-formatted design journal accompanied by the student instruction manual. Students were directed to work on their own designs, but they were encouraged to discuss them with their classmates. While the students were iterating, the teacher was instructed to circulate the classroom, check on student progress, and answer questions.

After students had had a chance to create at least two to three designs, they were introduced to the AITA. Using a genetic algorithm with preset parameters, the AITA used the current student design as the starting point, generated new designs by mutating the current design, improved its strategy by learning from the analysis result of each iteration, and evaluated a total of 50 new designs over five generations. While the AITA iterated through different designs, the students could view an animation of how one design changed into the next design and evolved into the final design over time. The final best design (including the design variables and the performance) was reported as an interactive graph alongside all previous iterations (Figure 2). Students then answered a series of reflection questions on a worksheet to document their reaction to the AI's design for use as feedback, i.e., AITA feedback, to think about how they could further improve their design. They then had until the end of the implementation to keep iterating either by interacting purely with AITA and/or with teacher feedback (Figure 3). During the implementation, students from both schools had access to three types of feedback—AI, teacher, and peers. The AI's design also doubled as a formative assessment of student design, because it was not guaranteed that the AITA could find a better design. If the student design was already close to the local optimum, then the AITA would be less likely to find a better design or improve by any significant amount. A post-survey was administered on the last day of the implementation.

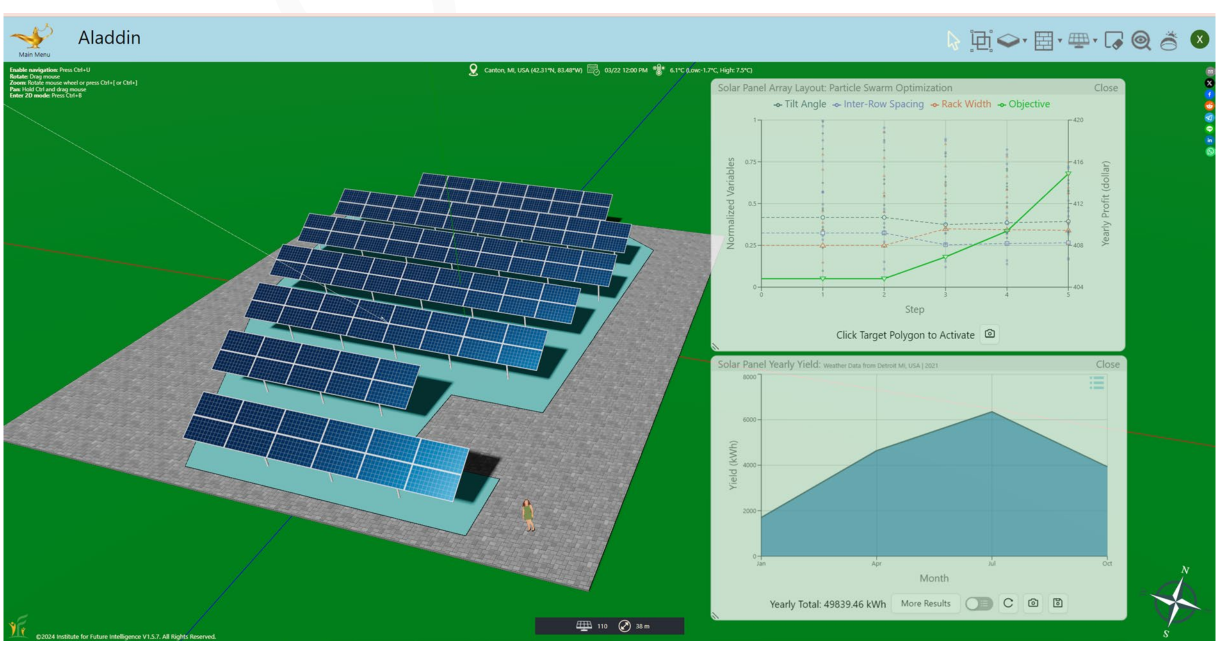


Figure 2. A sample screenshot of an AI-generated solar farm model in Aladdin. The top window shows the evolution of three design variables (tilt, RW, and IRS) and one objective (yearly profit) over multiple iterations. The bottom window shows the yearly yield analysis of the current design.

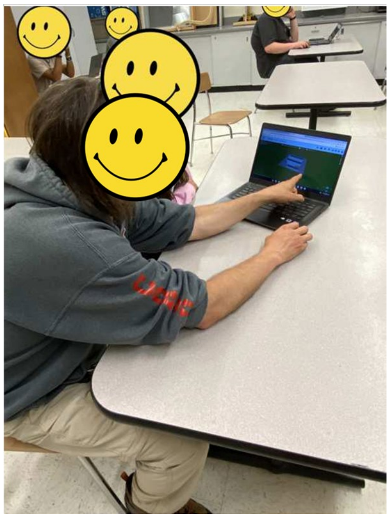

**Figure 3.** Classroom photos taken during the implementation. The science teacher from School 2 was giving feedback on one student's solar farm design. Three students on the back were inputting design variables into Aladdin's layout wizard to test their solar farm designs.

**Table 2.** A list of interview questions related to design feedback.

| Relevant interview questions | Question Count |
|---|---|
| What part of the design process was the MOST engaging for you? | 13 |
| What part of the design process was the LEAST engaging for you? | 13 |
| What kind of feedback did you receive on your design, if any? | 12 |
| How would you compare receiving design feedback from a teacher, a classmate, and AI? | 7 |
| If you can change one thing about how AI gives you design feedback, what would you change? Why? | 7 |
| If you can receive feedback on anything, what kind of feedback do you think will help you the most with improving your design? | 10 |

## Evaluation

The goals of the evaluation were to address the following research questions (RQs): 1) To what extent and in what ways does the Solar Farm Design curriculum affect students' achievement of learning outcomes? 2) To what extent and in what ways does the use of AITAs affect students' achievement of learning outcomes? 3) To what extent and in what ways does the use of AITAs affect students' perception of AI?

### *Data collection*

The main data were collected using pre- and post-surveys, worksheets, and interviews. Students filled out the same survey before and after the implementation of the Solar Farm Design project to determine their pre- and post-activity performance. The survey consisted of the following components: a) Multiple choice questions that assess student understanding of the following solar energy science concepts (Xie et al., 2018; 2023): i) *Daily changes of solar angles*, ii) *Seasonal changes of solar angles*, iii) *Projection effect*, and iv) *Optimal solar panel tilt angle*; b) A two-tier question to elicit students' knowledge about engineering design and its processes; c) An engineering design self-efficacy survey consisting of nine Likert-type items with a possible 100 points (adapted from Carberry et al., 2010); d) Likert scale questions about student perception of the AITA after design activity (adapted from Kim et al., 2020); e) Two open-ended questions about what the students enjoyed and would have changed about the curriculum (note that components 'd' and 'e' were only included in the post-survey).

After receiving AITA feedback, students were asked to answer a series of reflection questions on their AI worksheets and student instruction (see sample questions on the "*Design Challenge (AI Worksheet)* tab on Supplement Document 1" and questions on pages 12–18 on Supplement Document 2) and given another opportunity to improve their solar farm design. They also recorded their pre-AI, AI, and post-AI designs on the AI worksheet as design documentation.

Three interviewers (AB, IL, & RJ) conducted 15- to 20-minute semi-structured interviews after the project implementation. Teachers selected five students from School 1 and 10 from School 2, based on availability and interest, covering different levels of engagement and performance. The lead interviewer was determined by availability. All three interviewers followed the same interview protocol developed by one researcher (RJ) (See Supplement Document 3), but the exact questions asked varied for each student based on their progress and time availability. Table 2 shows some relevant interview questions that were centered around how students experienced different types of feedback.

In addition, the following types of supporting data were collected and used to corroborate the main data:

- Design artifacts: During the Solar Farm Design project, students were instructed to save their solar farm models on the Aladdin cloud storage. When present, these files were used to validate the design documentation on the AI worksheets.
- Student activity log data: Every student action in Aladdin (such as turning on the heliodon, changing the tilt angle of a solar panel, and simulating the yearly energy output) was automatically logged and stored in a database. This log data was consulted when the design documentations on the AI worksheets were incomplete or contained inconsistencies. Note that the data logger in Aladdin was only enabled during the implementation for research purposes, and it is currently disabled for regular users.
- Design journals: During the solar farm design portion of the project, students were asked to document each design iteration in a design journal and answer a series of reflection questions about each iteration (see "*Design Challenge Journal*" tab on Supplement Doc2). The journals were also used to validate the design documentation on the AI worksheets.
- Teacher feedback: After each class period, the teachers in both schools gave verbal descriptions of their observations in the classroom. An informal interview was also conducted after the project implementation.
- Observation notes: In both schools, the teachers set up additional cameras that were connected to a video conference during the implementation, so researchers

could observe and take notes on the classroom dynamics and student engagement.

Compliant with IRB requirements, student assent forms and parent consent forms were distributed prior to the study, and data were only collected from assenting students. Each student was assigned an anonymous ID in the following format: C_P1S1 or M_P1S1. The prefix indicates the school (C for School 1, M for School 2), and P1S1 stands for "period 1, student 1." All students were only referred to by their anonymous IDs in subsequent data analysis.

## Data analysis

After omitting students without consent forms or incomplete data, the remaining student data that were available for analysis are presented in Table 3. Unless otherwise stated, the pre- and post-survey data referenced below include only data from 80 students in School 1 and 28 students in School 2 who submitted both surveys. A data table showing the alignment of available data with the corresponding research question(s) and relevant learning objectives is summarized in Table 4. To answer RQ1 relating to the effect of the engineering design activity on students' learning outcomes, we conducted a paired one-tailed $t$-test using the aggregate results from the pre- and post-surveys.

For RQ2 concerning the impact of AITAs on students' learning outcomes on the understanding of engineering design processes and engineering design performance, we coded students' responses to the two-tier question on the pre- and post-surveys: (1) "*How familiar are you with 'engineering design'*? (multiple choice question)", and (2) "*What are the important components of the engineering design process? Name at least three components*. (open-ended question)" Closed coding of the responses (Saldaña, 2021) was done by RJ. Eight engineering process components, including *identify a design need, research a design need, develop design solutions, select the best possible design, construct a prototype, evaluate and test a design, communicate a design*, and *redesign* (e.g., Carberry et al., 2010; Massachusetts Department of Education, 2006) were adopted to code the open-ended responses.

We also analyzed participants' design data to investigate the effect of AITA on students' engineering design performance in response to RQ2. Students' design data (including the pre-AI design, AI design, and post-AI design) were organized from AI worksheets, which were corroborated with their design journals and validated using the log data. A student's design data was considered to be "complete" if it contained all three designs. A student's design data was considered to be "coherent" if there were no inconsistencies among different data sources. A student's design data was considered to be "unique" if the student did not share their data with anyone else in a group setting. Forty of 111 students in School 1 left complete, coherent, and unique documentation of their solar farm designs before receiving AI

Table 3. A breakdown of student data availability from different sources.

| | School 1 | | School 2 | |
|---|---|---|---|---|
| Pre-survey | 97 / 111 | 87.39% | 31 / 31 | 100.00% |
| Post-survey | 87 / 111 | 78.38% | 28 / 31 | 90.32% |
| Both pre- and post-surveys | 80 / 111 | 72.07% | 28 / 31 | 90.32% |
| Design journal | 40 / 111 | 36.04% | 16 / 31 | 51.61% |
| AI worksheet | 24 / 111 | 21.62% | 1 / 31 | 3.23% |
| Interview | 5 / 111 | 4.50% | 10 / 31 | 32.26% |

Table 4. A data table showing the alignment of available data with the corresponding research question(s) and relevant learning objectives.

| Data Collected | Data Analyzed for Corresponding RQ(s)[a] | Relevant Learning Objectives[b] |
|---|---|---|
| ***Both Pre- and Post-Survey*** | | |
| Multiple choice questions that assess student understanding of the following solar energy science concepts: i) Daily changes of solar angles; ii) Seasonal changes of solar angles; iii) Projection effect, and iv) Optimal solar panel tilt angle | 1 | A1-A4 |
| A two-tier question to elicit students' knowledge about engineering design and its processes | 2 | B1-B5 |
| An engineering design self-efficacy survey consisting of nine Likert-type items with a possible 100 points | 3 | |
| ***Post-Survey Only*** | | |
| Likert scale questions about student perception of the AITA after design activity | 3 | B1-B5 |
| Two open-ended questions about what the students enjoyed and would have changed about the curriculum | 3 | B1-B5 |
| ***AI Worksheet*** | | |
| Design documentation for pre-AITA, AITA, and post-AITA designs | 2 | B1-B5 |
| Reflection questions | 2 | B1-B5 |
| ***Student Interviews*** | | |
| Interview transcripts | 3 | B1-B5 |
| ***Supporting Data*** | | |
| Pre-AI design journals | 2 | B1-B5 |
| Solar farm design artifacts stored on the cloud | 2 | B1-B5 |
| Student activity log data | 2 | B1-B5 |
| Student instruction manual (some reflection questions) | 3 | A1-A4,B1-B5 |
| Teacher feedback | 2,3 | B1-B5 |
| Observation notes | 2,3 | B1-B5 |

[a]Goals of Evaluation (Research Questions): To what extent and in what ways does… RQ1) the Solar Farm Design curriculum affect students' learning outcomes? RQ2) the use of AITAs affect students' learning outcomes? RQ3) the use of AITAs affect students' perception of AI?

[b]Learning Objectives: *A) Solar Energy Science exercise*: 1) Describe the Sun's position using solar elevation and azimuth angles; 2) Describe the daily and seasonal changes of solar angles; 3) Describe the relationship between the angle of incidence and the energy output of a solar panel; 4) Explain the optimal tilt angles of a solar panel that maximizes the energy output in each season and in a year. *B) Solar Farm Design unit*:1) Evaluate a design solution using the given criteria and constraints; 2) Collect evidence of the design performance using computer simulations; 3) Improve the design performance through iterations; 4) Create a design that meets the given criteria and constraints; 5) Explain the choice of design variables using scientific principles.

feedback on the AI worksheets, which were considered to be valid data to be included for further analysis henceforth. Of the 71 students who were excluded from further discussions of design performance, three students worked together with other students and used their data with permission; 15 documented data that were incomplete or incomparable with other students' data; 27 used other students' data without express permission or documented data that could not be validated by other data sources; and 26 did not finish the activity or document enough coherent data. Of the 40 students in School 1 with valid data on AI worksheets, 24 documented the feedback from AI; 10 were from a class that had to end early before the AI activity; and six did not document enough data. Of the 24 students who received AI feedback, 14 increased the yearly profit of their final design; two did not find a better design than AI's recommendation; three did not iterate again or document enough data; and five students already had near-optimal designs (see Figure 4a). Since the AITA was unlikely to improve a near-optimal design within one run (or 50 iterations), students with near-optimal designs were excluded from any analysis or discussion of AI feedback. We only consider the 16 students who either improved in their final design or accepted the AITA feedback for further analysis.

As for School 2, 16 of 31 students left complete, coherent, and unique data of their solar farm design before receiving AI feedback; seven documented data that were incomplete or incomparable with other students' data; and eight did not finish the activity or document enough data. Of the 16 students with valid data, 10 documented the feedback from AI; six did not document enough data. Of the 10 students who received AI feedback, only one student increased the yearly profit of their final design; nine did not iterate again or document enough data (Figure 4b). Due to the lack of recorded design improvement taken from the AITA feedback from School 2, only AITA feedback for 16 students in School 1 was manually categorized by RJ by comparing AI's designs with students' pre-AI designs.

Students' engineering design performance was evaluated using a single metric: The yearly profit of their solar farm design, which equals the revenue (determined by the total energy output of the solar panels) minus the cost (determined by the number of solar panels used). In general, a tilt angle equal to the latitude of the location (around 42° for the two schools) would optimize the energy output per solar panel for an entire year, although the curriculum placed a wind resistance constraint that limited the maximum tilt angle to 35°. The other two design variables, row width and inter-row spacing, were coupled: A larger RW required a larger IRS to avoid inter-row shading. Therefore, an optimal design was one with an optimal tilt angle and a suitable pairing of RW and IRS that fit as many solar panels onto the given plot as possible while minimizing inter-row shading. Solar farm designs would be hereinafter denoted in the following format: (tilt, RW, IRS).

The allowed design space was specified as (0°-35°, 1–6 panels, $2\,m \sim 10\,\mathrm{m}$), which was also set as the search range of the AITA's genetic algorithm. At least three local optima existed within this design space: (35°,1 panel, 2.3 m), (35°, 2 panels, 4 m), (35°, 3 panels, 7 m). When deployed on the given plot in the curriculum, these optimal designs produced a yearly profit of around \$420 in School 1 and \$517 in School 2. The difference was due to different weather conditions. A student design was considered to be optimal or near optimal if its yearly profit was greater than \$400 in School 1 or \$500 in School 2.

To better illustrate the common themes in AI feedback and student reactions in response to RQ3 about the effects of AITAs on students' perception of the usefulness of AITA feedback and self-efficacy, we analyzed students' AITA perception responses on the post-survey and engineering design self-efficacy (measured by confidence) survey (Carberry et al., 2010) by averaging students' pre- and post-score on each item. We also analyzed students' different reactions toward AITA feedback and compared their post-AI design. In addition, interviews were transcribed by three researchers (AB, IL, & RJ) and analyzed through inductive thematic analysis (Braun & Clarke, 2006) by RJ. The initial open codes were generated after a thorough reading of the transcribed student responses to all interview questions. In a second round of focused coding, only responses containing open codes related to AI or feedback were reviewed, and

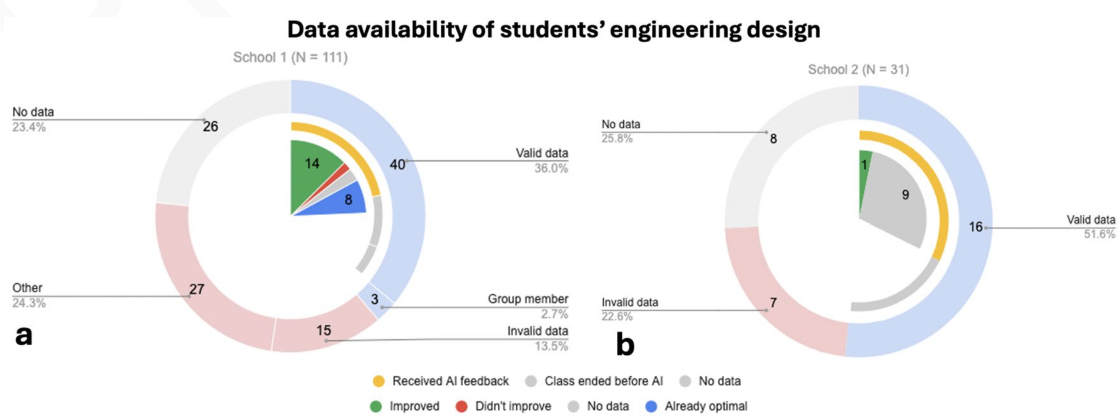


**Figure 4.** A breakdown of the data availability of students' engineering design. The outer ring showed the sample size of student designs before the AI feedback and reasons why data points were excluded. The middle ring showed the sample size of students who received AI feedback and reasons for exclusion. The inner pie chart showed students' reactions to AI feedback. (a) Statistics from School 1. (b) Statistics from School 2.

emergent themes were identified from the final codes and refined into sub-themes.

## Results

### Solar energy science knowledge learning outcome

To answer whether the Solar Farm Design curriculum affects participants' conceptual understanding (RQ1), we analyzed the items related to solar energy science knowledge on pre- and post-surveys. In School 1, the mean score increased in all four questions that assess solar energy science knowledge (Figure 5). In School 2, the mean score increased only in question 3 (Figure 5c), remained the same in question 1 (Figure 5a), and decreased in questions 2 and 4 (Figures 5b, d). In multiple choice questions Q1 and Q3, School 1 outperformed School 2 in both the pre- and post-surveys by a margin of 10%-20%. In Q2 and Q4, where School 2 outperformed School 1 in the pretest, School 2 actually performed worse in the post-survey, whereas School 1 still improved. The paired one-tailed *t*-test indicated that students at School 1 improved their understanding of solar energy science after the Solar Farm Design project ($p_1$ = .00002, $M_{pre\text{-}1}$ = 1.15, $SD_{pre\text{-}1}$ = 0.9; $M_{post\text{-}1}$ = 1.8, $SD_{post\text{-}1}$ = 1.08), but students at School 2 did not ($p_2$ = .25, $M_{pre\text{-}2}$ = 1.04, $SD_{pre\text{-}2}$ = 0.74; $M_{post\text{-}2}$ = 0.89, $SD_{post\text{-}2}$ = 0.74).

### Engineering design performance and engineering design process learning outcome

To determine the effects of AITAs on participants' engineering design performance and engineering design process learning outcome (RQ2), we summarized the results from relevant pre- and post-survey questions and their post-AI designs in the following sections.

### Familiarity with engineering design

Before the project, when prompted with the question about how familiar they were with engineering design on the pre-survey, 51.5% of the students in School 1 and 51.6% in School 2 chose level 2 (i.e., "*have some idea what it is, but don't know when or how to do it*"), followed by level 1 or the lowest level ("*I have never heard of it or I have heard of them but don't know what it is.*") with 35.1% and 35.5% of the votes, respectively (See Supplement Figure 1a,b). In the post-survey, over half of the students (54.0%) from School 1 reported level 4 or the highest level (i.e., "*I can explain what it is, how to do it, and I have done it*") after the project, followed by level 2 (29.9%) (Supplement Figure 1c). In School 2, the plurality choice remained at level 2 (42.9%), while 21.4% of the students selected level 4 (Supplement Figure 1d).

### Understanding of engineering design process

When asked to identify important components of the engineering design process on the pre- and post-surveys, students frequently mentioned more "hands-on" processes like "develop design solutions" and "construct a prototype," a consistent observation across both surveys that reflected the common impressions of engineering design among these students (see Figure 6). They rarely mentioned processes like "identify a design need," "select the best possible design," and "communicate a design," which was expected

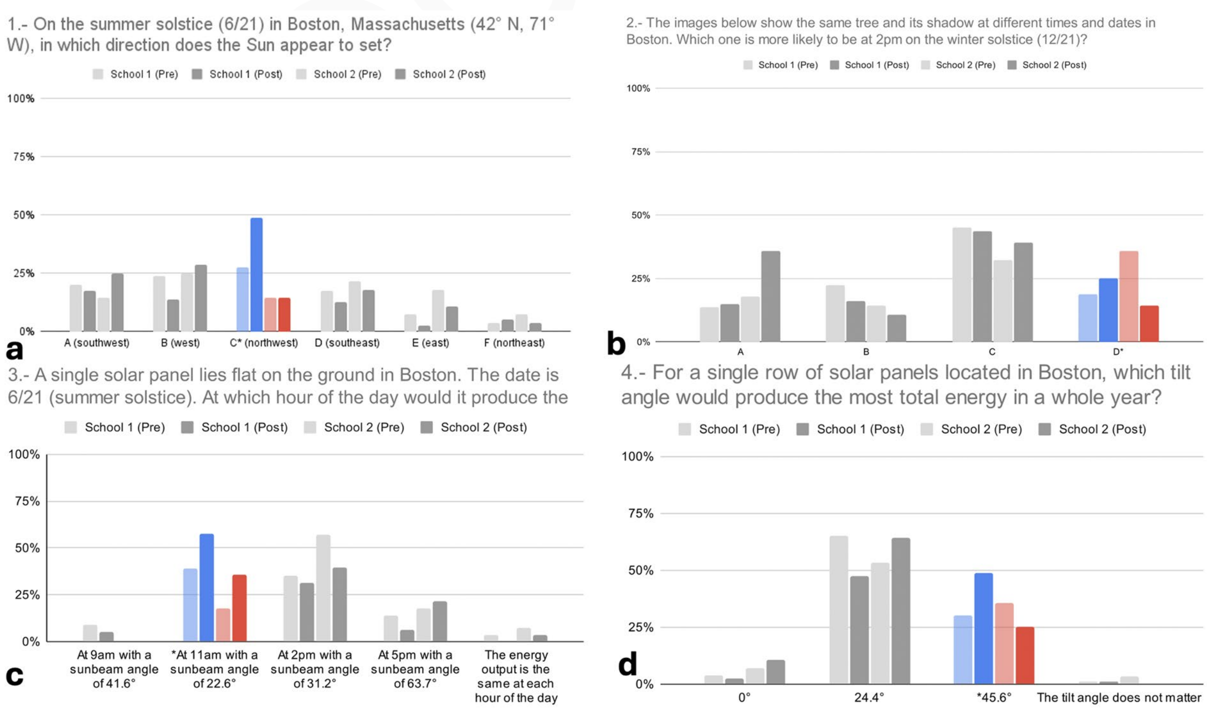


**Figure 5.** The distribution of student responses to the solar energy science assessment items, before and after completing the solar farm design curriculum. The correct answer is marked with an asterisk symbol. (a) A question about daily changes of solar angles. (b) A question about seasonal changes of solar angles. (c) A question about the projection effect. (d) A question about the optimal solar panel tilt angle.

because those were not emphasized or practiced heavily in the project. Students in School 1 mentioned "research a design need" and "evaluate and test a design" more frequently than those in School 2 in the pre-survey, but the differences were reduced in the post-survey. This suggested that the Solar Farm Design project helped students develop a more holistic impression of the engineering design process and reduced the achievement gap between different student populations. Finally, both schools saw significant increases in the popularity of "redesign" (including responses that mentioned "iterate", "multiple designs", or "trial and error"), which may in part be due to the interaction with the AITA and the extensive redesigning that many students went through.

Students' understanding of the engineering design process shows that the percentage of non-responses, which included answers like "*I do not know*" and "*I am not sure*," decreased by 23% in School 1 and 25% in School 2, respectively (Figure 6). For six of the nine categories, the differences between pre- and post-surveys were within 10%. In School 1, 25% more students identified "redesign" as an important component of the engineering design process. In School 2, the biggest increase was for "research a design need" (+15%), followed by "redesign" and "evaluate" (+10%). Students from both schools reported similar levels of familiarity with engineering design in the pre-survey. While 54% of students from School 1 reported the highest familiarity with engineering design in the post-survey, only 21.4% did so in School 2.

## *Design space and benchmark performance*

Figure 7a shows that in School 1, 36 out of 40 students were able to make a profit on their solar farm design before receiving AI feedback ($median_{Profit\text{-}PreAI\text{-}1}$ = \$282, $SD_{Profit\text{-}PreAI\text{-}1}$ = 362.40). Sixteen of the 36 students received AI feedback that improved the yearly profit of their designs ($median_{Profit\text{-}AI\text{-}1}$ = \$392, $SD_{Profit\text{-}AI\text{-}1}$ = 28.92). Fourteen students further improved their profit after receiving AI feedback ($median_{Profit\text{-}PostAI\text{-}1}$ = \$420, $SD_{Profit\text{-}PostAI\text{-}1}$ = 32.50), whereas two students did not. Notice that the students who made lower profits before receiving AI feedback were the ones who showed the most improvement with their post-AI designs.

Figure 7b shows that in School 2, 11 out of 16 students were able to make a profit on their solar farm design before receiving AI feedback ($median_{Profit\text{-}PreAI\text{-}2}$ = \$266, $SD_{Profit\text{-}PreAI\text{-}2}$ = 561.00). Only one student recorded the AI feedback, which failed to improve the yearly profit due to a technical issue in the early version of the Aladdin software. However, the

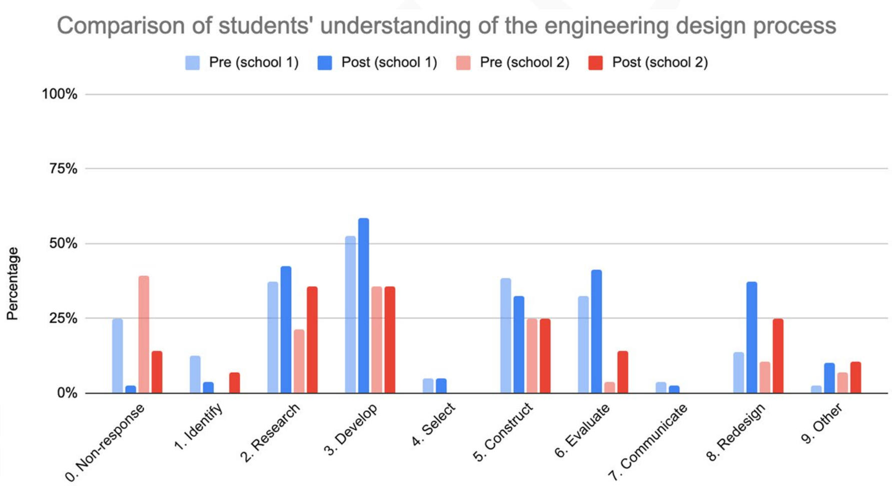


**Figure 6.** A comparison of students' self-reported understanding of the engineering design process, before and after completing the solar farm design curriculum calculated from semi-structured interview codes.

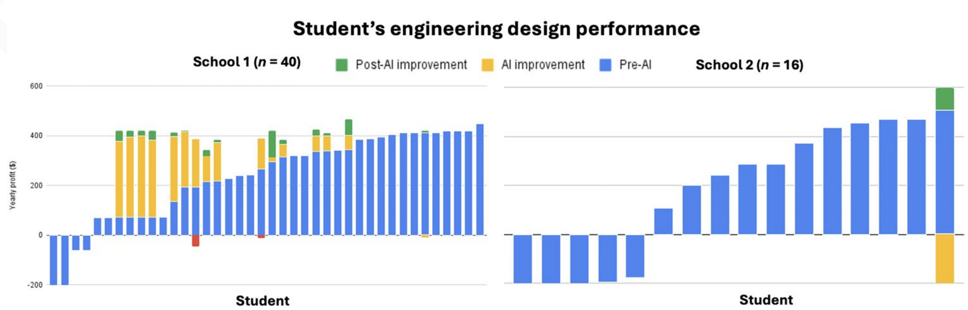


**Figure 7.** A comparison of students' engineering design performance (measured by the yearly profit of their final solar farm design), before receiving AI feedback (blue), the improvement made by AI's recommended design (yellow), and the improvement made by students' design after they received AI feedback (green). The negative y-axis was truncated to save space. Red columns below zero indicated the difference in yearly profit between students' post-AI design and AI's design when a student couldn't further improve AI's design. Yellow columns below zero indicated the difference between students' pre-AI design and AI's design, when the AI couldn't improve a student's design, which were rare faulty behaviors from AI. (a) Statistics from School 1. (b) Statistics from School 2.

student improved their own design after receiving AI feedback, nonetheless.

### Students' reactions to AITAs

In the following sections, we combined the results from surveys, AI worksheets, reflection questions, observations, and interviews to address RQ3 regarding AITA's impacts on participants' perceived usefulness, self-efficacy, design decisions made based on AI feedback, and reactions and surprises toward AITA.

#### Impact of AITA on perceived usefulness and self-efficacy

Figures 8 and 9 demonstrate the overall impact of implementing AITA feedback on learning engineering design based on responses from the survey. Students' self-rated confidence with the eight components of the engineering design process indicates that the average ratings were 45 (School 1) and 36 (School 2) out of a possible 100 before the project and increased to 66 (School 1) and 58 (School 2) after the project (Figure 8). However, students from School 2 reported confidence levels that were consistently 5–15% lower than students from School 1 did across all engineering design processes. Similarly, even though both schools saw a roughly 20% overall increase in confidence levels in the post-survey, students from School 1 reported higher confidence levels for every design process. Similarly, students' overall perception of AITA's usefulness in the post-survey shows that students were more positive in School 1 with an average rating of 4.38 compared to 4.25 for School 2 ($M_1$ = 4.38, $M_2$ = 4.25), although students' perceptions displayed a larger standard deviation for School 2 ($SD_1$ = 1.53, $SD_2$ = 1.90) (Figure 9).

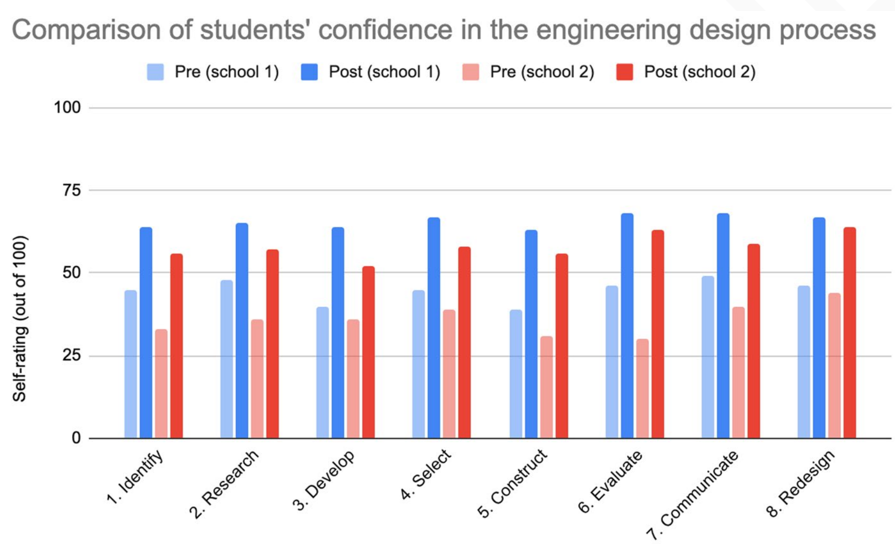


**Figure 8.** A comparison of students' self-rated confidence in the engineering design process, before (pre-) and after (post-) completing the solar farm design curriculum using a 100-point Likert-type survey.

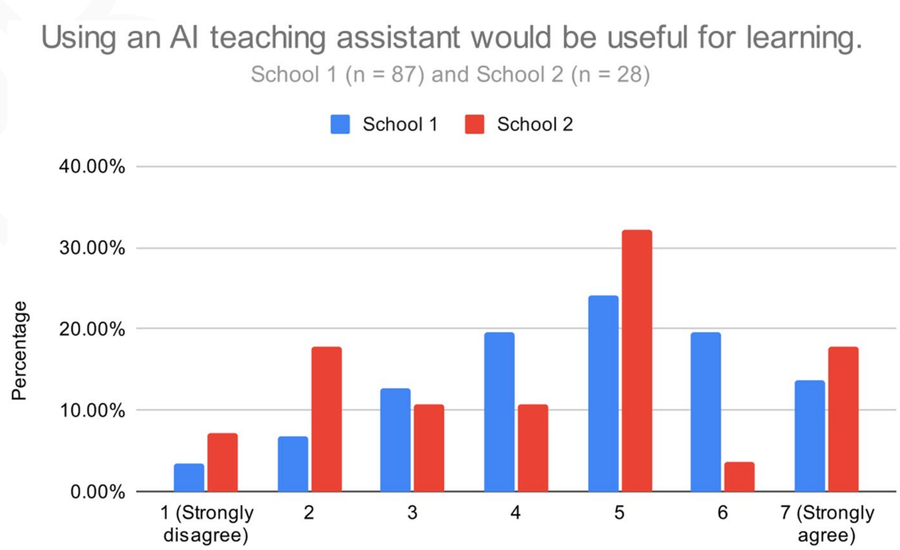


**Figure 9.** Responses to a 7-point bipolar Likert scale question about student perception of the AITA, with 1 being "strongly disagree" and 7 being "strongly agree." The remaining numbers were not labeled in the survey.

### Summary of students' reactions to their pre- and post-AI design

Students' reactions to the AI feedback from the AI worksheet and reflection questions are summarized in a vertical treemap chart (see Figure 10). The tilt angle was analyzed separately due to it being relatively independent from the other two design variables. One common AI feedback to decrease the tilt angle, received by 11 students total, was considered an artifact of the software algorithm because the AI search range for tilt angle was capped at 35°, and AI could only output floating point numbers that were bound to be less than 35°. Ten of 11 students made the correct decision to reject the feedback to decrease the tilt angle. When AI did use a higher tilt angle, which was interpreted as feedback to optimize the tilt angle, two students improved their design by accepting the feedback and further increasing the tilt angle, while one accepted the feedback but did not make improvements in their post-AI design. We also found that when AI used design variables that were similar to what students used, the feedback became more difficult to interpret and led to some misinterpretation. For instance, one rare occasion showed that AI was able to inspire students to diverge the tilt angle even when its own attempt yielded less profit (see case C_P5S16's reaction in Category 1 below for a more detailed example).

Another common AI feedback was to use a different RW or IRS than the student did, and it was interpreted as the feedback to diverge more in the design space. There were 14 students who received this feedback, and among them, 12 made further improvements in their post-AI designs, i.e., nine copied AI's exact same values in their final designs and three accepted the feedback and diverged further, while the other two did not lead to any improvement (one accepted and one rejected AI feedback). We found two incidences in which students were inspired by AITA's recommendation to optimize IRS. Even though one accepted and one rejected AI's feedback, they both improved their post-AI designs. One student (see C_P5S16's case for example below) was inspired by AITA's suggestion and diverged for 17 unsuccessful iterations. He then copied AITA's RW and IRS design in his post-AI design and further improved the annual profit.

### Categorizing students' reactions to AI feedback

Based on students' reactions to AITA's recommendations on the three design variables from AI worksheets and reflection questions, the 16 students' reactions are grouped into five categories:

#### Category 1: Best of both worlds (seven students)

The first category illustrates the most common AI feedback and the most common student reaction. When AI parameters yield better outcomes, the students accepted or copied those parts. When students' outcomes were better, they rejected AI-recommended parameters. For instance, C_P5S29's pre-AI design was (35°, 2 panels, 12 m) and made a yearly profit of $220. AI's recommended design was (33.87°, 3 panels, 8 m) and made $374 of profit. The student was "not really" surprised by AI's tilt angle because they "*knew the optimal range is between 30°-60°*", but they rejected AI's change and kept the original tilt angle in the final design.

However, C_P5S29 was "very much" surprised by the other changes because "*minor changes created such drastic differences.*" They ended up copying AI's exact RW and IRS in the final design, (35°, 3 panels, 8 m), which made $383 of profit.

C_P5S16's pre-AI design was (35°, 2 panels, 4 m) and made $412 of profit. When they saw AI's design of (33.21°, 3 panels, 6.89 m), they were not surprised by the wider rows, keeping the following notes in the AI worksheet: "*In my*

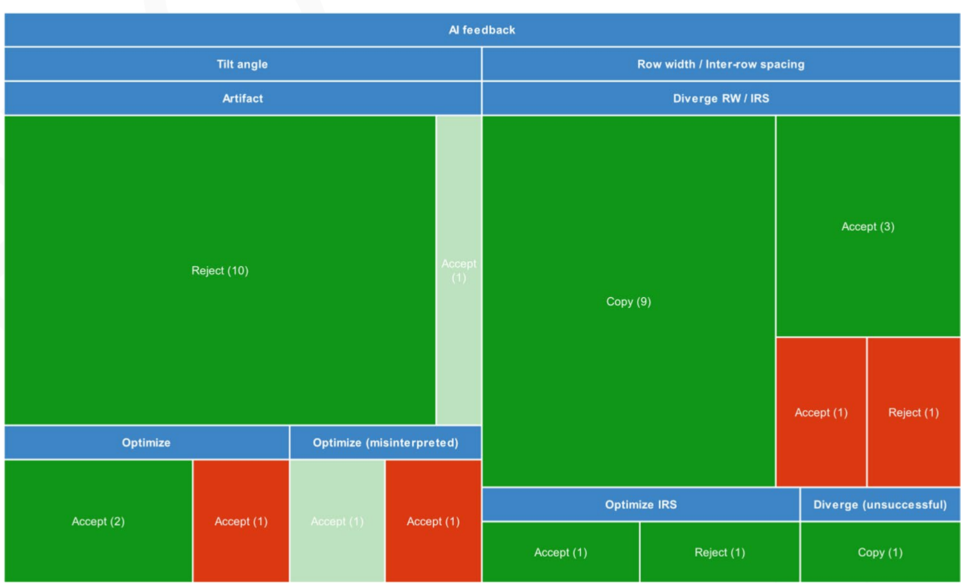


**Figure 10.** A vertical treemap chart showing students' reactions to different AI feedback. The chart can be interpreted using the following color-coded tiles—Blue tiles represented different categories of AI feedback. Dark green tiles represented student reactions that led to further improvement of the yearly profit in the post-AI design. Red tiles represented reactions that didn't lead to any improvement. Light green tiles represented reactions that wouldn't have led to any improvement per se but still did due to the optimization of other design variables. The number included in the parenthesis represents the number of students for a particular category.

*initial testing, I found that 3 panel width was effective in creating a large amount of profit. Only after much more testing did I find a much more effective 2 panel width design. So, it does not surprise me that the AI found and stuck with the 3-panel width idea.*" Their design documentation on the AI worksheet, which had over 10 iterations, showed that they had already tried (35°, 3 panels, 6.5 m). Their post-AI design of (35°, 3 panels, 6.89 m) combined AI's wider rows and their original tilt angle and made $421 of profit.

In total, seven students responded in the same way to AI feedback: They kept their original tilt angle when AI used a lower one, copied AI's RW and IRS, which were different from their original ones, and improved their profit in the end by combining the best of both worlds.

### Category 2: Go the extra mile (three students)

Instead of directly copying AI's design variables, students in Category 2 accepted AI's feedback to diverge and diverged further in the design space. In the case of C_P4S30 (see Table 5), the student's pre-AI design was (35°, 2 panels, 7 m) and made a yearly profit of $297. The AITA recommended the design (33.87°, 3 panels, 8 m) and increased the yearly profit to $314.44. The student also reported being surprised by AI's changes to the RW because "*an increase in the [row] width by 1 panel resulted in such a big difference in energy output.*" They went on to explain the rationale of this surprise: "*If there are more solar panels per row, then the yearly energy output will increase because the additional panels receiving sunlight contribute to the amount absorbed.*"

In the reflection, the student agreed with the change in RW, while acknowledging that there could be more optimal tilt angles. He did not agree with the change in IRS, citing the decrease in solar panels that would ensue. The student then documented three design attempts (Figure 11), the best of which made $420.62 of profit by integrating AI's new RW of three panels into their original design to create the final design (35°, 3 panels, 7 m). C_P4S30 ended up accepting AI's feedback to diverge and diverged further in the design space by combining AI's wider rows with their own tighter IRS.

Two other students reacted similarly by going the extra mile. For instance, C_P6S18's pre-AI design was (32°, 2 panels, 6 m), and when they saw AI's design of (30.66°, 3 panels, 7.37 m), they pivoted to wider rows and further increased the IRS, creating the final design (35°, 3 panels, 8 m). However, C_P6S20's pre-AI design was (35°, 3 panels, 10 m), and when they saw AI's design of (33.76°, 1 panel, 2.33 m), they kept their wider rows but took inspiration to decrease the IRS, creating the final design (33.76°, 3 panels, 6 m).

### Category 3: You may already know this (two students)

Students in this category were able to activate their existing knowledge based on AI's feedback to improve their design and reinforce science and engineering design understanding. For example, when AI changed the tilt angle of C_P4S19's design from 30° to 34.5°, they further increased it to 35°.

**Table 5.** A partial reproduction of C_P4S30's AI feedback worksheet showed the documented evolution of their solar farm design, before and after receiving feedback from the AITA.

| Design variables | Your design | AI's final design | Your new design 1 | Your new design 2 | Your new design 3 |
|---|---|---|---|---|---|
| Tilt angle (°) | 35 | 27.47 | 35 | 30 | 30 |
| Row width (panels) | 2 | 3 | 3 | 3 | 2 |
| Inter-row spacing (m) | 7 | 7.55 | 7 | 8 | 5 |
| Number of solar panels | 66 | 96 | 99 | 90 | 90 |
| Yearly energy output (kWh) | 30094.77 | 43305.87 | 45044.48 | 40803.92 | 40790.26 |
| Yearly Profit ($) | 296.69 | 314.47 | 420.62 | 345.98 | 342.57 |

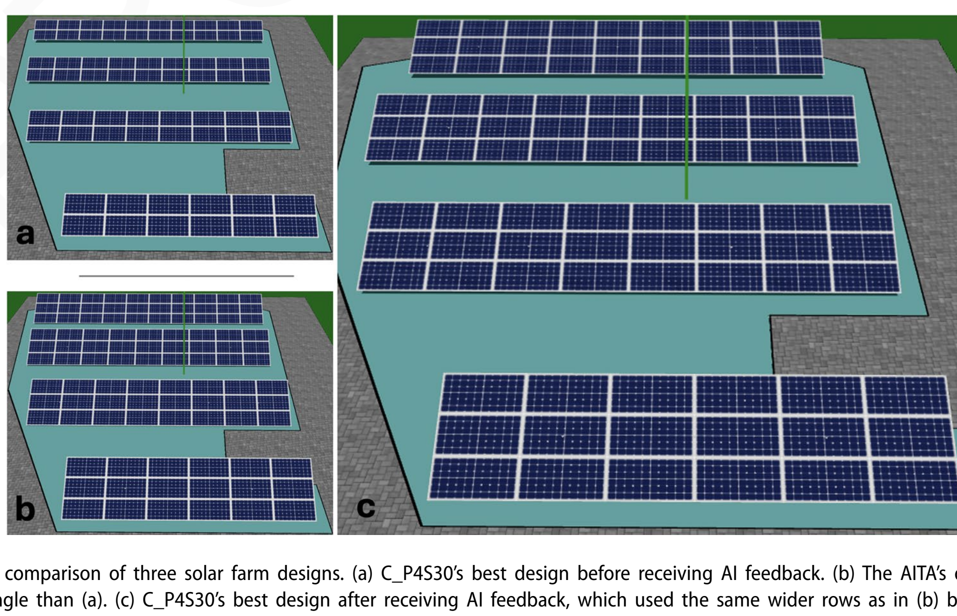


**Figure 11.** Side-by-side comparison of three solar farm designs. (a) C_P4S30's best design before receiving AI feedback. (b) The AITA's design, which used wider rows and a lower tilt angle than (a). (c) C_P4S30's best design after receiving AI feedback, which used the same wider rows as in (b) but reverted to the higher tilt angle used in (a).

Reflecting on this decision, they wrote: "*We noted that any degree closer to 35 is more ideal. So even that slight change of half a degree can make a dollar or more difference over the course of a year.*" They also copied AI's RW and IRS to increase yearly profit by $4. They concluded by stating: "*Increasing tilt closer to the 35 degree constraint will result in an increase in profit. The act of using less solar panels far [outweighs] the costs of spending … on [maintaining] more solar panels poorly positioned.*"

Another student (C_P5S23) was surprised to see the AITA increase the tilt angle of their design from 30° to 33° and wrote: "*I say this because I did not think about further tilting the panels as I felt I had reached the sweet spot. Now, I understand that the further tilt accommodates the middle seasons, fall and spring, much more appropriately and therefore produces more energy.*"

### Category 4: Unintended consequences (two students)

Students in this category accepted but misinterpreted AI's feedback, which led to unexpected results in their final design. For example, when C_P5S5 compared AI's tilt angle of 30.09° with their own tilt angle of 30°, he wrote: "*I wasn't surprised because I feel like 30 degrees is the best tilt angle.*" They did try other design variables after receiving AI feedback with varying design variables, (31°, 4 panels, 9 m), (35°, 3 panels, 10 m), (27°, 2 panels, 4 m), but when none of the attempts improved AI's design, they concluded that "*the [30°] angle is perfect.*" Similarly, when C_P5S20 compared AI's feedback of (32.74°, 1 panel, 2.85 m) with their own iterations (35°, 1 panel, 3 m), (31°, 1 panel, 2.5 m) they wrote: "*I am not surprised because I know that the tilt angle is most efficient in the low thirties due to the seasons.*" When their best post-AI design, which used a slightly higher tilt angle of 31° and a narrower IRS, turned out to be better than AI's, they concluded that: "*If the solar panel tilt is closer to 30 degrees, the amount of total energy increases*" without further iteration by controlling variables. Both C_P5S5 and C_P5S20 interpreted AI's feedback based on their feelings or relied primarily on their prior knowledge rather than conducting more controlled experiments to isolate the effect of different design variables on annual profit.

### Category 5: No pain, no gain (two students)

Students in this category had the potential to improve their design just by changing a particular variable at a time based on AI's recommended design for a few more iterations. For example, C_P5S2's pre-AI design was (30°, 1 panel, 3 m), and when AI recommended the design (32°, 3 panels, 7 m), they rejected the feedback to diverge to 3-panel rows and kept using 1-panel rows. That decision could still have led to success, since the student was already near the local optimum (35°,1 panel, 2.3 m), which would have made over $400 of profit, but the student stopped short at (34°,1 panel, 2.5 m), which made $10.5 less than AI did. On the contrary, C_P5S1 demonstrated strong reasoning skills (by changing one design variable recommended by AITA at a time) and persistence (with 50 iterations recorded on log data) by making the most profit ($466) among their peers.

### Students' preferred source of feedback and their reasons

Students described various qualities they valued when receiving feedback on their designs. Table 6 shows the results of the thematic analysis of the interview transcripts and Figure 12 and Figure 13 show students' evaluation about the AITA curriculum. In general, AI feedback excelled at providing a visual representation of iterative computation and epistemic agency with this computational power (Figure 12). For example, when asked from whom they prefer to receive feedback, C_P5S29 picked AI feedback because AI "*can do more calculations than the teacher can or would want to.*" Nonetheless, AI lacked the empathy that a human teacher or a peer could provide. When M_P5S16 explained why they preferred peer feedback, they said, "*…because they [my classmates] are in the same kind of position as me. Because they're also experimenting with it as we do it together, and it kind of helps. Sometimes we might realize something and be like, 'Oh, yeah, make sure to do that.'*" There were also four accounts of students complaining about the software crashing or being difficult to navigate, which hindered its usefulness, and could have been attributed to the high frequency of complaints around technical issues (Figure 13).

Students reacted favorably to the type of comparative feedback provided by the AITA. M_P6S9 found the AI feedback "straightforward" because "*it was like seeing, 'Oh, I did this one way, and the AI did this another way*'." Students also offered a variety of ideas for how the AITA could give different types of feedback. M_P5S16 would like more specific feedback, such as "*it's a little bit too slightly to the left.*" C_P3S28, who did not interact with AI, also wanted directive feedback such as "*Oh, other students had success changing the row width. Maybe you should try doing that.*"

In addition to design improvements, some students reported that AI helped them figure out what to do in the design process. For example, M_P5S10 described how the AITA was "*where it all clicked, and it made sense on what everything was doing.*" He explained that

> …before we started making our own [solar farm], I was just kind of pressing buttons and watching things change… I didn't know really what to do… And then, once we did our own [solar farm], the directions actually made me change them and I watched them change, like the visual aspect.

Finally, there was further evidence that the presence of an AITA created affective responses among students. Students reported being surprised by the AITA, which was also observed in the student reflections from the AI worksheet. M_P5S10 also claimed the most engaging part of the design process was "*when … we built one [solar farm], and then we compared [ours] to the AI.*" The engagement did not guarantee a positive experience throughout, but there could be payoff at the end. According to M_P5S5, "*I was hearing how a little bit of people were struggling, but then they were happy when they beat the AI.*"

Student interview data showed no dominant source of feedback (teacher, peer, or AI) that was preferred by the majority (see Supplement Figure 2). Instead, students provided different reasons for preferring each source of feedback. For example, M_P6S7 elaborated that they "*might take in my fellow classmates if they were doing really good, and*

**Table 6.** A list of themes and subthemes from the interviews.

| Themes | Sub-themes | Definition | Example quote | Count |
|---|---|---|---|---|
| Values in feedback | Visualization | The feedback recipient can visualize designs and design changes in different representations. | C_P5S25: "I'm a super visual learner. So, being able to see the whole program work itself through and see the pictures behind it and be like 'oh, you can see it's at this angle' or 'this is where the shading part is' and 'this is the best part to put this.' That's what I really like about it." | 2 |
| | Authority | The feedback giver is deemed trustworthy due to their perceived experience in the subject matter. | C_P5S25: "The teacher won't know everything. The teacher can take a guess, but AI probably knows the best design because you can see it run through 100 designs." | 2 |
| | Empathy | The feedback giver can express emotions and relate to shared experiences. | M_P5S3: "Mr. [teacher] isn't a robot. He can show emotion, so he was smiling, where[as] the bot was just giving me information of what I could do to do better." | 2 |
| | Usability | The feedback recipient can receive, understand, and act on the feedback easily. | C_P4S19: "It [AI] was like a lot of things that you would press and steps you gotta go through. That was confusing." | 4 |
| Types of feedback | Comparative | The feedback shows the recipient other people's results. | M_P5S10: "It was nice to actually compare [to] someone's numbers." | 5 |
| | Directive | The feedback provides explicit steps the recipient needs to take to improve. | M_P5S16: "If you move it at a slanted angle and push it backwards a little bit, you might have better results." | 4 |
| | Facilitative | The feedback engages the recipient in independent thinking and sensemaking. | M_P5S10: "Maybe it [could] explain why they thought more panels were better, or why [changing] 10 to 12 makes a huge difference." | 2 |
| Affect from feedback | Surprise | The feedback leads the recipient to unexpected conclusions. | M_P6S9: "I was surprised by it, because obviously [it was] a lot better than mine." | 1 |
| | Challenge | The feedback recipient views the feedback as a challenge to win a competition. | M_P5S10: "We need to try and beat it because it made it a challenge."' | 3 |
| Additional effects of Feedback | Intention | The feedback helps the recipient figure out the goals and next steps. | M_P6S9: "[I] kinda just… thought, 'you know I'll just put in some random number and see what comes out'. But as I [saw AI's design, I] realized that maybe I didn't need to have something lower or something higher… The AI knew that right off the bat." | 3 |
| | Confirmation | The feedback confirms whether certain decisions are good/bad for the recipient. | C_P5S29: "I just learned that my assumptions about the tilt angle were correct." | 1 |

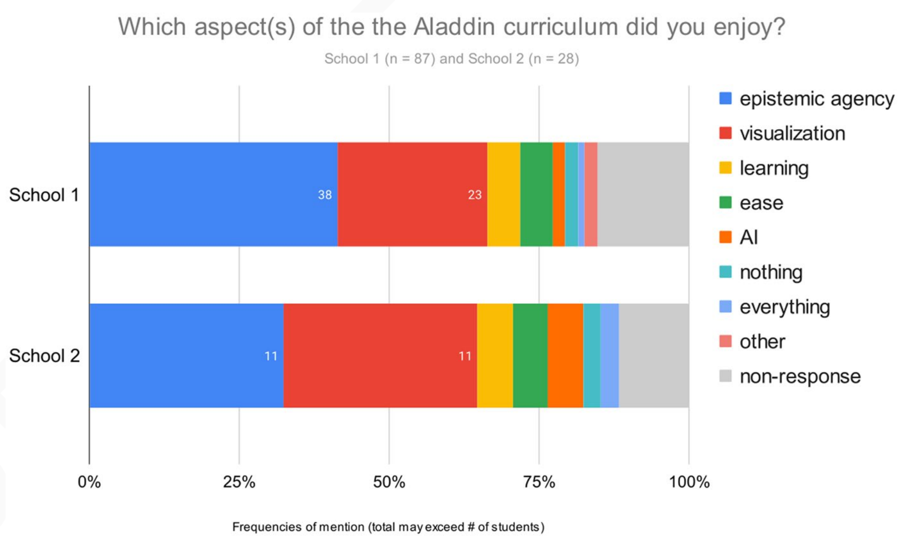


**Figure 12.** A breakdown of the student responses to the open-ended question in the post-survey about what they liked about the curriculum.

*they were improving multiple times each time.*" It is worth noting that some students also expressed a preference for multiple sources of feedback. For example, M_P5S16 commented "*I would compare it [the feedback] and see if a lot of the things that they mentioned matched up with one another… If a student and a teacher said that the positioning was a little bit awkward, because it's 2 [people], then it has more of a stronger standpoint.*" It was also observed that students frequently (though not always) discuss with their peers during their independent work time.

The teacher from School 2 found the AITA to be useful in an informal conversation after the implementation, explaining that "*in a lot of cases, the kids just need a suggestion.*" He also suggested other types of feedback that the AITA could give, from more directive feedback ("*You have tried to change the tilt angle 5 times, how is it going for you so far? You have not touched the other variables*") to more facilitative feedback ("*I see your number went down. Here are some reasons it may have gone down instead of up as you expected.*").

## Student evaluation

In the free-response questions about the learning experience in the post-survey, students left both praises and criticisms

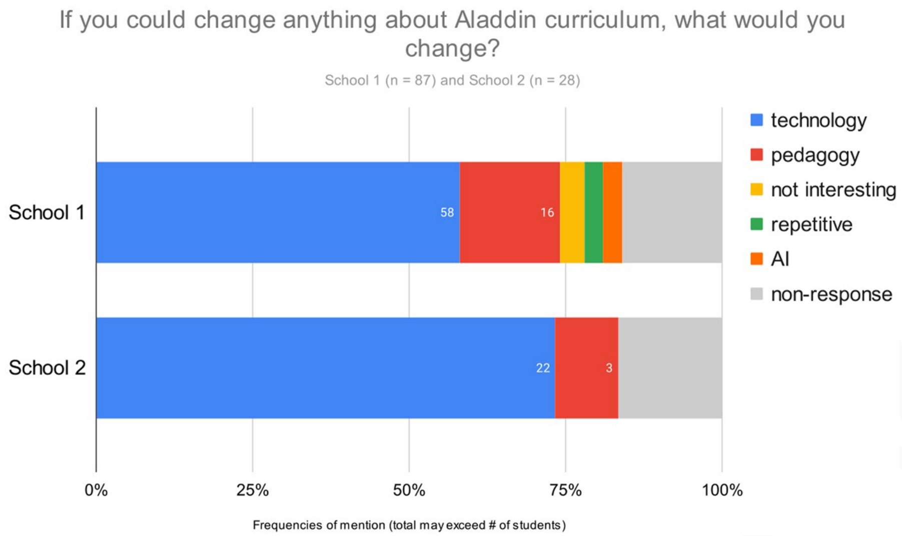


**Figure 13.** A breakdown of the student responses to the open-ended question in the post-survey about what they would change about the curriculum.

of the Aladdin curriculum (Figures 12 and 13). On the positive side, 49 out of the 115 students surveyed mentioned the freedom and autonomy to design and experiment (Figure 12). For example, C_P6S2 stated: "*I liked playing around with different factors and seeing how they affected the outcome, whether that was the revenue or the cost or the amount of kWh that were produced as I changed different things.*" C_P6S28 was one of the 34 students who commented positively on the visual aspect of the Aladdin curriculum, stating that "*once I saw [AITA's] take on how to do it, it pointed me in the right direction to then make more profit and energy, which lead [sic] me to one of my final designs.*" Four students mentioned the AITA specifically, claiming that they enjoyed "*the use of the AI to find the best design for the solar panel farm*" (C_P4S16) because they "*liked seeing the AI work and go through many [iterations] quickly*" (C_P6S5).

However, 80 of the 115 students experienced some level of technological difficulty, reporting issues with freezing, lagging, or navigation. Students also provided constructive criticism of the current implementation of AITAs (Figure 13). One student wished that "*when the AI says the [design] parameters are off, it should explain how they are off. I should not have to figure it out*" (C_P5S2). Another student wished to "*really see the math behind the AI and have an explanation [of] why and how AI works*" (C_P4S18). Finally, one student cautioned that "*using an AI as a teaching method should be used sparingly and a human teacher should be used a majority of the time*" (C_P4S33), with no specific explanation of their reasoning.

## Discussion

### *Science and engineering learning outcomes in response to RQ1*

Understanding solar science concepts in depth helped students consider different variables and perspectives that could facilitate and expedite their search for the optimal design logically. Results from the multiple-choice questions in the pre- and post-surveys suggested that students achieved some improved understanding of solar energy science concepts after finishing the Solar Farm Design project. Student performance differed between the two schools. The difference was most noticeable in two of the four questions (Q2: *seasonal change of solar angles*, and Q4: *optimal solar panel tilt angle*), where the correct answers did not receive the majority vote in the post-survey. Option C in Q2 was the most popular distractor as opposed to the Key in Option D. The only difference between these options is the length of the shadow. The shadow in Option C is shorter, which is a logical, educational guess for the solar angle two hours past noon. Students were able to rule out the shadow pointing to the north (Option A) and no shadow (Option B), but not the "height" of the sun during different seasons. Even though students could observe the seasonal change using the heliodon tool (see Figure 1), they might not have attended to different effects caused by the seasonal change of solar angles. In Q4, the most popular distractor students opted for was approximately half (24.4°) of the most optimal tilt angle (~42°), which should reflect the corresponding latitude of the place, i.e., 42° N. The finding is indicative that many students were either not transferring or generalizing the rule from the AITA activity to solve a similar problem. Notice that we put a cap of 35° on the tilt angle in our design due to the wind resistance constraint; therefore, students might not be able to extrapolate from the limited range of tilt angle to answer Q4.

### *Advantages of AITAs in enhancing science and engineering learning outcomes in response to RQ2*

A comparison of students' design performance before and after using the AITA, their design reflections, and their interviews provided both quantitative and qualitative data that indicated the pedagogical advantages of using AITAs to provide feedback on student learning.

#### *Playing the role of peer assistants*

The use of AITAs is comparable to learning from peers (Wood & O'Malley, 1996). This was found to be a common theme in the interviews, with at least five out of 15 students identifying a need or preference for comparing their results with those of someone else, be it another student or an AITA. In an in-person learning environment, students may spontaneously compare their design performance (yearly profit) with others. While this peer feedback is not directly actionable without further comparison of design variables and processes, it informs them of their relative standing in the group and what has been proven to be achievable in terms of design optimization. Similarly, the AITA provides a concrete design with quantitative performance data as a comparison, but it also exceeds typical peer feedback in two aspects: 1) While students sometimes only share their design performance but not their design variables, the AITAs always provide both pieces of information, so that students can quantitatively compare both designs, identify specific changes that contribute to the improvement (if any), and act upon it; 2) While students rarely discuss their thought process or design rationale that led to their designs, the AITA visualizes the design process in an animation, where students can see the evolution of one design iteration to the next. The abstract concept of divergence and convergence cycles can also be visualized: Students can see that the earlier design iterations look more variable, and the later iterations look less variable. This visual aspect facilitates learning from contrasting cases (Schwartz et al., 2016) and may be responsible for helping novice students adapt their design strategy from random trial and error to a more systematic approach. Such visual benefit may explain why learners perceived visualization effects as the second most enjoyable factor of the AITA curriculum.

### Playing the role of experts

The use of AITAs is also comparable to learning from experts. Assessing large numbers of engineering design solutions has always been a difficult task, especially for geoscience educators who may not have equal expertise in engineering design. The AITA provides an efficient approach for formative assessment: If the AI design performs much better than the student design, then it obviously means that the student design has lots of room for improvement. However, if the AI design does not show much improvement, then it means that the student design may already be close to the optimal solution (see Figures 7 and 8). AI-generated feedback also exceeds expert feedback in two aspects: 1) While expert feedback is often based on design heuristics that still need to be tested, AI-generated feedback is supported by quantitative evidence. 2) Student interview data suggests that the dynamic visualization of the computation process increases student confidence in the authority of the feedback, which may encourage the adoption of such feedback.

### Creating cognitive dissonance with surprising design

Finally, there is evidence to support the hypothesis that the psychological effect of AITAs, especially the element of surprise, contributes to student learning. Students in Category 2 documented how AI's surprising design prompted them to apply previously acquired knowledge about solar energy science to explain the surprise and identify ways to further improve the design. C_P4S14, one of the students in Category 3, explained how AI's surprising design triggered an important understanding of solar energy science: Seasonal changes of solar angles are an important factor in choosing an optimal tilt angle. These findings are consistent with the theory of cognitive development and constructivist learning (Lutz & Huitt, 2004): When presented with new knowledge that does not fit into any existing schema, the cognitive dissonance may trigger students to restructure their existing schema to *accommodate* the new knowledge.

### Activating and reinforcing existing knowledge

An analysis of the different types of AI feedback and student reactions showed that one of the most prominent effects of the AITA in Aladdin was either reinforcing or activating students' existing knowledge. For the 10 of 16 students that correctly rejected AI feedback to decrease the tilt angle, their reflections showed that many already knew the optimal range of the tilt angle and therefore raised doubts when the AI feedback contradicted the solar energy science knowledge they had already learned. As for three of the 16 students accepting AI feedback to increase the tilt angle, their reflections also showed that they were already aware of the optimal range of tilt angle, but for whatever reason, they did not fully optimize it (35°) on their initial attempt. In this case, the AI feedback activated their existing knowledge and reminded them to optimize it further.

### Fostering more divergent thinking

Another effect of the AITA was allowing students to create more divergent designs. The effect was most prominent in the case of the three students who went the extra mile to find a new RW-IRS pairing for their post-AI design. For the eight students who simply adopted AI's exact RW-IRS pairing in their post-AI design, less could be said about whether the students learned anything about the coupled nature of RW and IRS, how they impacted design performance, and how divergent thinking could lead to better designs. However, it could be viewed as a case of collaborative intelligence between humans and AI (Wilson & Daugherty, 2018), where each contributed what they knew about the design problem and took advice from each other.

Nonetheless, C_P5S16 presented a rare case, where the AITA failed to improve the student's original design. Their reflection referred to the "initial testing" and "much more testing" that they had done to discover two of the local optima in the design space. Their activity log data also confirmed that they had already iterated 30 times before receiving AI feedback and another 17 times afterward, which might explain why the student still managed to learn even from a failed attempt from AI: The student was very engaged and already explored the design space pretty thoroughly, giving them an edge in distilling helpful information from the raw feedback.

On the other hand, two students did not further improve their designs after receiving AI feedback, likely due to not having iterated enough times. Two other students reinforced their premature conclusions about the tilt angle despite having improved their designs, likely due to changing multiple variables at a time instead of running controlled experiments. Observations from these cases suggested that the quality and extent of the effect of the AITA varied based on many factors such as the students' level of engagement/persistence, existing knowledge, familiarity with engineering design practices, etc., and that additional scaffolding may be necessary to support the learning in one or more of those aspects.

### Student perceptions of AITAs in response to RQ3

Student perceptions of AITAs are mixed. In general, the AITA was regarded as being able to meet some students' needs but not all. Based on a combination of data sources, such as surveys, observations, and interviews, students in School 1 rated the AITA more favorably than those in School 2, which may be partially attributed to the fact that the students in School 2 exhibited more overall disengagement with the AITA curriculum. These mixed responses agreed with a previous study that found variation in student perceptions of automated feedback (Calvo & Ellis, 2010). Student interview data further supports the positioning of AITAs not as a substitute for teacher or peer feedback but as an additional source of feedback that can amplify the effect of human feedback (when available).

Results from the self-assessments also indicate that students achieved increased self-efficacy and enhanced their epistemic agency (Miller et al., 2018) in engineering design after the project. The difference in engagement in AITA activities was also salient, considering that only one out of 31 students documented valid AI data. Out of all students who submitted both the pre- and post-surveys and had the opportunity to interact with AI in class, 12.5% (10 of 80) students from School 1 did not document anything, while 64.2% (18 of 28) students from School 2 did not document anything. These differences in self-efficacy and engagement may be inherent, given that the teacher from School 2 had commented on the behavioral issues in his classes, and there had been multiple accounts of observations where the teacher had to address such issues publicly in class. Nevertheless, improvements could be made to support the struggling students and address disengagement.

### Room for improvement

Data from this preliminary study of AITAs in geoscience and engineering education generated important insight into how this pedagogical approach can be improved, specifically in the following aspects:

#### More informative feedback

Many students thought that the AITA could give "more specific" feedback. Recommendations collected from the students and teacher can be categorized as (1) directive feedback, which includes pointing out what the student did wrong and telling the student explicitly what they need to do to improve the result; (2) facilitative feedback, which includes explaining the reason why the student did not perform as expected, asking the student to reflect on patterns in their process or behavior, or recommending that the student experiment with something new. It is worth noting that students and the teacher preferred different types of feedback. Students generally wished for immediate actionable feedback that quickly improved their design product and required less cognitive effort. In contrast, the teacher recommended feedback that focuses more on scientific reasoning than immediate action, targets the design process or mindset more than the product, and requires more cognitive effort. Because students exhibit diverse needs that vary greatly depending on their levels of prior knowledge and current progress (Schwartz et al., 2016), it seems more desirable for AITAs to provide multiple types of feedback rather than canonizing any one approach.

#### More psychological support

Student interview data also highlighted the importance of human interaction in the classroom. While the current AITA lacks the emotional attention of a teacher and the shared experience of a fellow student and makes clear that future research should explore how AI can support students in the psychological dimension in addition to the cognitive dimension, we would like to reiterate that the intention of this innovation is not to create any replacement for human interaction. Rather, it is to create an alternative to accommodate students' diverse needs and a fallback in times of disruption.

#### More inclusive and equitable AI

Our findings resonated with other studies with AI feedback (e.g., Shi & Aryadoust, 2024) that students who were more engaged in the project benefited more readily from the current implementation of AITA, while students who were disengaged appeared to have gained less from their interactions with the AITA. Taking students' engineering design performance from School 1 for example (see Figure 7a), students with less desirable initial designs benefited more with more AI engagement, and most of these students' final designs outperformed those who either did not adopt AI's suggestions or reached near optimal with initial attempts. To make AI more inclusive and equitable, future research should focus on improving the usability of the software to enhance motivation and boost engagement by providing both directive and facilitative AI feedback that addresses different student needs. To this end, large language models (LLMs) may be well suited for the task of providing automated and personalized feedback (Shi & Aryadoust, 2024) in educational contexts to support more inclusive learning (Chen et al., 2023).

## Limitations

There are several limitations to this evaluation. The first limitation is that more validated assessment items specific to solar energy engineering were not readily available; otherwise, we could adopt multiple assessment items to elicit students' ESS knowledge of items like Q2 and Q4. Second, it was difficult to assess and control for students' prior knowledge of solar energy or engineering design during the study, so some patterns in student responses to the AITA may be attributed to students' prior knowledge. For example, students who had more familiarity with the subject matter may be more engaged in reflecting on the feedback they received and therefore exhibit more learning gains and a more positive attitude toward the AITA.

In terms of data analysis, the discussion of students' engineering design performance was based on a comparison of three snapshots within the entire design trajectory, namely the pre-AI, AI, and post-AI designs, instead of the evolution of all iterations. The potential correlations between factors such as engagement and performance were not discussed. In addition, due to the solo coding and analysis of qualitative data from one researcher, the inter-rater reliability was not measured. These limitations could also provide directions for future research.

## Implications

The study has exposed long-standing shortcomings in geoscience education, such as inadequate integration with engineering design at the K-12 level, and the lack of alternative feedback mechanisms. However, with a wealth of experience accumulated with a call for integrated STEM education (NGSS Lead States, 2013), geoscience educators are now in a much better position to proactively build a more equitable and resilient learning environment that cultivates student agency concerning global challenges.

Part of the solution entails deeper integration of engineering design into geoscience education, which would afford students the opportunity to apply their science knowledge to solve pressing problems of today and tomorrow. ESS educators interested in using the Solar Farm Design curriculum are welcome to explore the full problem space of solar energy engineering, a booming industry in demand of a greater workforce. For example, the profitability of the same solar farm design varies greatly in different geographic regions and depends on factors such as the weather and the local electric rate. A profitable solar farm design in the Northeastern US may not be profitable at all in the Midwestern US due to fewer sunshine hours and lower electric rates, which can lead to rich discussions about the relationship among geoscience, engineering design, and public policy. In addition, Aladdin supports the use of custom ground images, which educators can use to overlay additional GIS data and discuss geological and environmental considerations during site assessment.

Similar to how online learning has transitioned from a novel concept to a common alternative during the unprecedented pandemic, another part of the solution is to introduce alternative forms of support into the learning environment, including the selective use of AI. While the AITA in Aladdin remains available 24/7, educators interested in Aladdin and its accompanying curriculum materials should be mindful of providing multiple feedback mechanisms to accommodate diverse student needs, with AI feedback being an additional option that complements existing methods. Future research may focus on making AI-generated feedback more understandable, e.g., by incorporating strategies recommended in the engineering design coaching tool suggested in this study and from our previous work (e.g., Purzer et al., 2022). Another potential research direction is to extend the preliminary work on AI as instructional design agents with different personas (Schimpf et al., 2019) and create more humanized agents that can assess students' psychological states using their design activity data and providing socio-psychological interventions (Yeager & Walton, 2011), in addition to design feedback. As AI continues to gain traction in education, the findings about how AI-assisted coaching tools may impact students' conceptual and affective learning will only become more relevant for applying AI in ways that benefit all students.

## Acknowledgements

The authors are indebted to the assistance of science teachers DK and RC, who allowed us to conduct research in their classrooms and implement the curriculum. The authors would also like to thank Dr. Alex Barco and Isaac Lyss-Loren for their contributions to data collection and analysis. This work was supported by the National Science Foundation (NSF) under grant numbers 2105695, 2131097, and 2301164. 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.

## Disclosure statement

The authors report there are no competing interests to declare.

## ORCID

Shannon Sung http://orcid.org/0000-0001-5704-2920
Rundong Jiang http://orcid.org/0000-0003-0298-1298

## References

Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Weegar, R., Wu, Y., & Li, X. (2021). Explainable AI for data-driven feedback and intelligent action recommendations to support students self-regulation. *Frontiers in Artificial Intelligence*, *4*, 723447. https://doi.org/10.3389/frai.2021.723447

Al-Shaery, A. M., Khozium, M. O., Farooqi, N. S., Alshehri, S. S., & Al-Kawa, M. A. M. B. (2022). Problem solving in crowd management using heuristic approach. *IEEE Access*, *10*, 25422–25434. https://doi.org/10.1109/ACCESS.2022.3156008

An, Y., & Mindrila, D. (2020). Strategies and tools used for learner-centered instruction. *International Journal of Technology in Education and Science*, *4*(2), 133–143. https://doi.org/10.46328/ijtes.v4i2.74

Bergey, P. (2020). Generative well pattern design—Principles, implementation, and test on OLYMPUS challenge field development problem. *Computational Geosciences*, *24*(6), 2079–2094. https://doi.org/10.1007/s10596-019-09912-w

Bethune, K. (2020). Changing trends and rethinking geoscience education in the context of a global crisis. *Geoscience Canada*, *47*(4), 167–169. https://doi.org/10.12789/geocanj.2020.47.164

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Carberry, A. R., Lee, H.-S., & Ohland, M. W. (2010). Measuring engineering design self-efficacy. *Journal of Engineering Education*, *99*(1), 71–79. https://doi.org/10.1002/j.2168-9830.2010.tb01043.x

Calvo, R. A., & Ellis, R. A. (2010). Students' conceptions of tutor and automated feedback in professional writing. *Journal of Engineering Education*, *99*(4), 427–438. https://doi.org/10.1002/j.2168-9830.2010.tb01072.x

Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial intelligence (ai) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers*, *25*(1), 161–182. https://doi.org/10.1007/s10796-022-10291-4

Darvishi, A., Khosravi, H., Abdi, S., Sadiq, S., & Gašević, D. (2022 *Incorporating training, self-monitoring and AI-assistance to improve peer feedback quality* [Paper presentation]. Proceedings of the Ninth ACM Conference on Learning @ Scale (pp. 35–47). https://doi.org/10.1145/3491140.3528265

Goldin, I., Narciss, S., Foltz, P., & Bauer, M. (2017). New directions in formative feedback in interactive learning environments. *International Journal of Artificial Intelligence in Education*, *27*(3), 385–392. https://doi.org/10.1007/s40593-016-0135-7

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *3*(1), 20–29. https://doi.org/10.1111/j.1745-6916.2008.00058.x

Haupt, G. (2015). Learning from experts: Fostering extended thinking in the early phases of the design process. *International Journal of Technology and Design Education*, *25*(4), 483–520. https://doi.org/10.1007/s10798-014-9295-7

Higgins, J. L. (1971). A new look at heuristic teaching. *The Mathematics Teacher*, *64*(6), 487–495. https://doi.org/10.5951/MT.64.6.0487

Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. *Mathematical Problems in Engineering*, *2022*, e5215722–19. https://doi.org/10.1155/2022/5215722

Hughes, B. B. (1974). Heuristic teaching in mathematics. *Educational Studies in Mathematics*, *5*(1), 291–299. https://doi.org/10.1007/BF01424551

Lucas, J. F. (1974). The teaching of heuristic problem-solving strategies in elementary calculus. *Journal for Research in Mathematics Education*, *5*(1), 36–46. https://doi.org/10.5951/jresematheduc.5.1.0036

Kim, J., Merrill, K., Xu, K., & Sellnow, D. D. (2020). My teacher is a machine: Understanding students' perceptions of AI teaching assistants in online education. *International Journal of Human–Computer Interaction*, *36*(20), 1902–1911. https://doi.org/10.1080/10447318.2020.1801227

Kishore, P., & Kisiel, J. (2013). Exploring High School Students' Perceptions of Solar Energy and Solar Cells. *International Journal of Environmental and Science Education*, *8*(3), 521–534.

Lutz, S., & Huitt, W. (2004). Connecting cognitive development and constructivism: Implications from theory for instruction and assessment. *Constructivism in the Human Sciences*, *9*(1), 67–90.

Massachusetts Department of Education (2006). Massachusetts science and technology/engineering curriculum framework. (Original work published 2001)

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053–1075. https://doi.org/10.1002/tea.21459

Mirchi, N., Bissonnette, V., Yilmaz, R., Ledwos, N., Winkler-Schwartz, A., & Maestro, R. F. D. (2020). The virtual operative assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PloS One*, *15*(2), e0229596. https://doi.org/10.1371/journal.pone.0229596

National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Committee on a conceptual framework for new K-12 science education standards. Board on science education, division of behavioral and social sciences and education*. The National Academies Press.

NGSS Lead States (2013). *Next generation science standards: For states, by states*. The National Academies Press.

Porter, B., & Grippa, F. (2020). A Platform for AI-enabled real-time feedback to promote digital collaboration. *Sustainability*, *12*(24) Article, 10243. https://doi.org/10.3390/su122410243

Porter, S. G., Greene, K., & Esposito, M. C. K. (2021). Access and inclusion of students with disabilities in virtual learning environments: Implications for post-pandemic teaching. *International Journal of Multicultural Education*, *23*(3)Article, 43–61. https://doi.org/10.18251/ijme.v23i3.3011

Purzer, S., Schimpf, C., Quintana-Cifuentes, J., Sereiviene, E., Lingam, I., & Jiang, R. (2022). Refine by design. *The Science Teacher*, *89*(6), 34–39. https://doi.org/10.1080/00368555.2022.12293711

Radford, J., Bosanquet, P., Webster, R., Blatchford, P., & Rubie-Davies, C. (2014). Fostering learner independence through heuristic scaffolding: A valuable role for teaching assistants. *International Journal of Educational Research*, *63*, 116–126. https://doi.org/10.1016/j.ijer.2013.02.010

Riggs, E. M. (2020). Access and success in geoscience education in the age of COVID-19. ED027–05.

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4E [Fourth edition]). SAGE.

Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. WW Norton & Company.

Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, *36*(2), 187–209. https://doi.org/10.1017/S0958344023000265

Shu, D., Cunningham, J., Stump, G., Miller, S. W., Yukish, M. A., Simpson, T. W., & Tucker, C. S. (2019). 3D Design using generative adversarial networks and physics-based validation. *Journal of Mechanical Design*, *142*(7) https://doi.org/10.1115/1.4045419

Sung, S., Jiang, R., Huang, X., & Xie, C. (2022). Panels v. Trees: Broadening the pathways of engineering education through integration with social studies. *Science Scope*, *45*(6), 24–33. https://doi.org/10.1080/08872376.2022.12291485

The United States Government (2021). Fact sheet: President Biden signs executive order catalyzing america's clean energy economy through Federal Sustainability. *The White House*, December 13. Retrieved August 5, 2022, from https://www.whitehouse.gov/briefing-room/statements-releases/2021/12/08/fact-sheet-president-biden-signs-executive-order-catalyzing-americas-clean-energy-economy-through-federal-sustainability/

Thornton, C. P., Ruble, K., & Jacobson, L. A. (2022). Education for children with chronic illness: Moving forward in online and virtual learning. *JAMA Pediatrics*, *176*(4), 341–342. https://doi.org/10.1001/jamapediatrics.2021.5643

U.S. Department of Energy (2021). *Solar futures study.* Retrieved August 5, 2022, from https://www.energy.gov/eere/solar/solar-futures-study

U.S. Energy Information Administration (2022). *Electric power monthly - U.S. Energy Information Administration (EIA)*. Electric Power Monthly. Retrieved August 5, 2022, from https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=table_es1a

van der Hoeven Kraft, K. J. (2017). Developing student interest: An overview of the research and implications for geoscience education research and teaching practice. *Journal of Geoscience Education*, *65*(4), 594–603. https://doi.org/10.5408/16-215.1

Vahdatikhaki, F., Salimzadeh, N., & Hammad, A. (2022). Optimization of PV modules layout on high-rise building skins using a BIM-based generative design approach. *Energy and Buildings*, *258*, 111787. https://doi.org/10.1016/j.enbuild.2021.111787

White, R., & Gunstone, R. (2014). *Probing understanding*. Routledge.

Wilson, H. J., & Daugherty, P. R. (2018). Collaborative intelligence: Humans and AI are joining forces. *Harvard Business Review*, *96*(4), 114–123.

Wood, D., & O'Malley, C. (1996). Collaborative learning between peers: An overview. *Educational Psychology in Practice*, *11*(4), 4–9. https://doi.org/10.1080/0266736960110402

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., Ding, X., Jiang, R., Santos, B. S., Magana, A. J., & Zara, J. (2023). Using computer graphics to make science visible in engineering education. *IEEE Computer Graphics and Applications*, *43*(5), 99–106. https://doi.org/10.1109/MCG.2023.3298386

Xing, W., Li, C., Chen, G., Huang, X., Chao, J., Massicotte, J., & Xie, C. (2021). Automatic assessment of students' engineering design performance using a bayesian network model. *Journal of Educational Computing Research*, *59*(2), 230–256. https://doi.org/10.1177/0735633120960422

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*(2), 267–301. https://doi.org/10.3102/0034654311405999

Yoo, S., Lee, S., Kim, S., Hwang, K. H., Park, J. H., & Kang, N. (2021). Integrating deep learning into CAD/CAE system: Generative design and evaluation of 3D conceptual wheel. *Structural and Multidisciplinary Optimization*, *64*(4), 2725–2747. https://doi.org/10.1007/s00158-021-02953-9