SUPPORTING AUTHENTIC PROBLEM-SOLVING THROUGH A CORNERSTONE DESIGN COURSE IN CHEMICAL ENGINEERING

ERIC BURKHOLDER, LISA HWANG, AND CARL WIEMAN Stanford University • Stanford, CA 94305

INTRODUCTION

D ngineers frequently engage with complex, ill-structured problems that require application of deep mathematical and scientific knowledge.^[1-5] Typically, engineering students encounter this kind of real-world problem-solving in the capstone design course.^[6] Over the past few decades, however, many programs have started to implement *cornerstone* design experiences, in which first-year engineering students are given the opportunity to engage in challenging design problems.^[7-8] Indeed, Ford et al. find that approximately 20% of introduction to engineering (or introduction to chemical engineering) courses have increased their design content in the past five years.^[9]

There are many models for cornerstone design courses described in the literature ^[10-16] and some models that demonstrate that first-year students can have meaningful experiences with real-world problems despite their limited technical background.^[17-19] These cornerstone design experiences have been shown to increase students' self-efficacy ^[20] and intellectual development^[21] and have also been shown to increase retention of women and underrepresented minorities.^[22] In the cognitive domain these design experiences have been shown to increase performance in future ill-structured problem-solving tasks.^[3,23]

Similar to the capstone design experience,^[6] the cornerstone design experience is widely considered to be successful. However, there are many different measures of success. For example, success may be measured through time spent engaged in cognitive tasks associated with design and problem-solving,^[10, 16-17] students' and instructors' thoughts on the impact of the design course,^[11, 15, 18] and increased retention.^[23] While all of these measures are important to the student experience in engineering education, we are primarily interested in the cognitive benefits of the cornerstone design course — in particular, problem-solving.

We could find no studies measuring the change in students' problem-solving skills over the course of the cornerstone design experience. Indeed, a more rigorous measure of success in terms of cognitive skills development would be to show that students' problem-solving skills become more expert-like during the cornerstone design experience.

Eric Burkholder, PhD, is a postdoctoral scholar and lecturer in physics at Stanford University. His research focuses on studying authentic problem-solving in undergraduate engineering programs, what factors impact student performance in introductory STEM courses, and mathematical sense-making in physics. He earned his PhD in chemical engineering from the California Institute of Technology.





Lisa Hwang, PhD, is a Senior Lecturer in the Chemical Engineering Department at Stanford University. She also serves as a Consultant with the Center for Teaching and Learning. She earned her M.S. and Ph.D. from Stanford University and her B.S. from MIT, all in Chemical Engineering. Before graduate school, she worked as a chemical engineer at Merck Research Labs in Rahway, NJ. She is interested in using research-based practices to improve the learning experience of her students.

Carl Wieman, PhD, is a professor of physics and education at Stanford University and the recipient of the 2001 Nobel Prize in physics. He is the founding director of PhET. His research focuses on authentic problem solving in STEM and medicine, student performance in introductory physics courses, and educational simulations.



© Copyright ChE Division of ASEE 2021

THEORETICAL FRAMEWORK

Problem-solving has been studied extensively in engineering education research,^[4-5, 24-27] cognitive psychology,^[28-30] and other fields of discipline-based education research (DBER).^[31-36] This research has revealed many key insights into human problem-solving and is quite diverse in scope.

Some research has focused on cognitive aspects of problem-solving, often using introductory physics problems as a context. This research has focused on problem-solvers' use of mental and physical representations,^[34] problem-solving strategies and knowledge structures,^[37-38] and the role of cognitive load and working memory.^[39-40] Other research has investigated expert-novice differences in problem solving to determine why experts are more successful. This research has looked at differences in procedures used by novices and experts ^[41-44] and differences in knowledge structures.^[28, 45-46]

There are several key limitations to research on problemsolving. First, the problems that researchers use to study problem-solving are typically limited in scope. Researchers in cognitive psychology have largely focused on knowledgelean tasks that discount the role of disciplinary knowledge in expert problem-solving.^[29-30] Researchers in DBER have largely used textbook-style problems (with some notable exceptions, e.g. ^[44, 47]) which do not necessarily represent the kinds of problem-solving that are used when solving workplace problems. We characterize these shortcomings as a lack of research on how people solve *authentic* problems.

Authentic problems require the application of deep disciplinary knowledge, are ill-structured,^[25] may require the problem-solver to collect additional information, and require the problem-solver to continuously reflect on their working solution and their problem-solving process.^[24] An example of an authentic problem would be an engineer designing a chemical process to produce a novel chemical product. The engineer has to determine the goals of the problem and how to achieve those goals. This kind of problem-solving is often iterative and requires the solver to reflect on their planning process and solution to reach a best solution for the problem.

Price et al. have developed an empirical framework of authentic problem-solving that is general across disciplines in science, engineering, and medicine.^[48] They conducted think-aloud interviews — adapted from the critical decision method of cognitive task analysis^[49] — with experts in a variety of disciplines to determine what are the decisions made by an expert as they solve an authentic problem. They found a set of 29 different decisions and 5 additional themes that were consistent across all disciplines.

Price et al. found that how experts make these different decisions and what qualifies as good decision-making is highly field-dependent and guided by a mental construct called a "predictive framework."^[48] A predictive framework

is a mental model of the key features of a problem and the relationships between those features. This framework allows the expert to explain observations and make predictions about how a system will behave (i.e. conduct mental simulations).

We assert that, to assess authentic problem-solving, one thus needs to develop a problem that requires students to make some of these 29 expert decisions and then evaluate how their decision-making compares to that of experts in the discipline. In previous work we described a general model for how to assess this kind of problem-solving [50] and have developed a specific example of this kind of assessment in the context of chemical process design.^[51-53] The basic structure of the assessment is to present the problem-solver with a non-functioning system, e.g. a flawed design, and then ask them general questions to see what features of the design they notice and what criteria they use to evaluate the design. We then ask increasingly more detailed questions about the design to capture a range of problem-solving skills. Experts and advanced students notice important features and flaws earlier in the assessment, whereas more novice student may never notice the flaws. We note that troubleshooting is a specific type of problem-solving that lends itself well to this type of assessment. As we detail below, it does not probe all problem-solving decisions, but does call for many of them and requires much of the same reasoning and knowledge. In this study, we use this assessment as a pre/post measure of students' problem-solving skills in a cornerstone design course to determine the cognitive benefits of such courses.

RESEARCH QUESTION

Our research question was: Does students' problem-solving improve during a cornerstone design course in chemical engineering? The purpose of this study is to determine the cognitive benefits of cornerstone design courses from a research perspective and to encourage educators to adopt a cornerstone design approach to teaching chemical engineering. This study provides evidence showing the cognitive benefits of the cornerstone design experience. In particular, it shows that students' authentic problem-solving improves during the cornerstone design course and suggests that these skills may improve because they are given opportunities to practice elements of authentic problem-solving.

COURSE CONTEXT

The subject of this study is a 2-credit, semester-long course in Chemical Process Design at a highly selective public research university in the western United States. The course is primarily taken by freshman and transfer students intending to major in chemical engineering. The course covers the basics of qualitative design, including separation and reaction operations. It also covers basic quantitative mass and energy balances, as well as economic analyses of chemical processes. Students' grades were based on 11 homework assignments (20%), a midterm exam (20%), a final exam (20%), a final individual project (20%), and participation (20%). Note that the researchers did not interfere with the design of this course; it was taught as it always has been by the instructor.

Class time is divided between a 50-minute weekly lecture and an 80-minute weekly recitation section. The lectures begin with a vignette about a notable chemical engineer to help students develop a sense of engineering identity. The majority of the lecture time is spent lecturing, but about 20% of the lecture time is devoted to targeted problem-solving practice. Students work in small groups to complete worksheets that are graded for participation credit or work individually to complete short quizzes that are graded for accuracy. The recitation sections (4 sections of 40 students) were led by a graduate student TA. Part of the recitation time was spent recapitulating concepts, terms, etc. from the lectures, but the majority of recitation time was dedicated to small group problem-solving. Students would work in ad hoc groups of 3-4 while the instructors went around the room to answer questions and check for understanding. The instructor would present the solutions to the recitation problems at the end of class, and students were allowed to ask further questions. This is the setting where students were able to practice decision making with timely feedback from instructors and fellow students, essential elements of deliberate practice.^[54] Outside of formal class time, instructors and TAs were available during scheduled office hours, as well as via an online forum where students could ask (and answer) questions, which was regularly monitored by the instructors.

The homework problems were mostly taken from Duncan and Reimer.[55] There was substantial variation in the types of problems that were assigned. The students completed some exercises that focused on "re-defining" the problem - reading a problem scenario and deciding what the true goal of the problem was. The students also completed some process analysis exercises in which they are given a flawed block flow diagram and asked to identify a certain number of errors and inefficiencies in the process. This uses many of the same decision-making processes as our assessment, but the textbook problems are more constrained and explicitly tell the students that there is something wrong with the process. Other problems were open-ended design tasks in which students were given information about the process chemistry and asked to design a process to produce a certain product (and later in the course to produce a product at a certain purity, or to optimize the profit made). These problems could require students to generate multiple potential solutions and reflect on how well these various solutions meet the problem goals. Other problems were well-structured exercises in mass and energy balances, in which the students made few decisions on their own. The exam and quiz problems were similar to the homework problems.

The individual final project in the Spring 2020 term was to design a system to produce food for a group of astronauts living on Mars. They could make a limited set of decisions, such as whether to use solar panels or nuclear power for the process or which crops to grow, and did some limited searching for information (e.g. food caloric densities). Students make a subset of the expert decisions during this project, some of which are probed by our external assessment (see Figure 1).

We administered an assessment of problem-solving (separate from other course assessments) as a pre- and post-test during the Spring 2020 semester. We describe the assessment in detail in the following section. The enrollment in the course was 140 students, and we collected 91 complete, matched pre- and post-responses that were open-ended. We chose a pre/post design to measure changes in problemsolving over the course, with the pre-test measurements indicating how much authentic problem-solving students had previously practiced. The assessment was assigned as part of the first and last homework assignment in the course, though students were only given credit for whether they made an earnest effort to complete the assessment (spent at least 10 minutes and submitted a complete response), not on the quality of their responses. Students were instructed to work independently, but we did not control the assessment conditions because it was impractical to take up class time with this assessment. In our analysis of written responses, we saw no evidence of collaboration.

ASSESSMENT

The chemical process design assessment, as well as the expert decisions that it probes, is shown in Figure 1. In the assessment the problem-solver is asked to take the perspective of a practicing engineer supervising an engineering intern. The intern develops a chemical process that has a number of flaws and inefficiencies. Diagnosing the flaws in this process requires some disciplinary knowledge but also requires the problem-solver to make a number of the expert decisions identified by Price et al.^[48] The assessment has four main parts. First, students are shown the nonfunctioning process and asked general questions about how they would evaluate the design. Second, students are shown a design that has the flaws corrected, but still is not optimized and asked more detailed questions about the design. Third, students are given a list of information relevant to the design and must decide how relevant the information is and how they can incorporate the information into the design. Finally, students are shown a more efficient process and asked whether it is a better design.

In a previous study ^[50] we collected responses to this assessment from three experts to validate the assessment and

Expert Problem-Solving Decisions

- · Decide upon the goals of the problem
- · Decide what are the important features
- · Decide which predictive frameworks are appropriate
- Decide how well the chosen solution holds
- Decide what approximations or simplifications are appropriate and reflect on assumptions and choice of predictive framework
- Decide whether additional information is needed
- Decide what are potential solutions
- Decide what is the best solution
- · Decide what calculations and data analysis are needed
- Decide how well does the chosen solution hold
- · Decide what are appropriate conclusions based on the data
- · Decide how well does the chosen solution hold
- Decide what are potential solutions
- · Decide what is the best solution
- · Decide what calculations and data analysis are needed
- · Decide whether the problem is solvable given the risks
- Decide how well does the chosen solution hold
- Decide what information is needed to solve the problem
- Decide how well the problem-solving approach is working
- Decide whether additional information is needed
- Decide upon priorities
- · Decide whether information is valid/reliable
- · Decide how information compares with expected results
- · Decide what conclusions are appropriate based on the data
- Decide how to follow up on anomalies
- Decide whether previous decisions about simplifications and predictive frameworks are appropriate
- · Decide what are potential solutions
- Decide what are appropriate conclusions based on the data
- Decide what is the best solution
- · Decide how well does the chosen solution hold
- · Decide what is the best solution

Assessment Question

Block flow diagram with errors and inefficiencies is now provided.

1. What criteria would you use to assess this design? Please list criteria in order of importance and explain your reasoning.

2. Given the information you have, does the design meet all your criteria? Please explain.

3. Is it possible to build a physically functioning chemical plant to produce tetrachloroethylene based on this design? Please explain.

4. [If no.] What modifications are necessary to make this process work physically?

Corrected but inefficient block flow diagram is now provided.

5. You notice some minor errors in the design and correct them; the corrected PFD is shown above. What other feedback would you give the intern regarding this design after making these corrections?

6. Given the information you have, is the design shown above optimal with respect to both material and energy consumption?

7. [If no.] What changes would you make to optimize the design with respect to material and energy consumption and why?

8. Are there safety concerns you have about this process in the event that it malfunctions? Please explain.

9. What additional information would you request from your intern to better evaluate this design and consider potential changes? Please explain in detail how you would use each piece of information.

List of information relevant to process operation, safety, and economics is now provided.

10. Your intern does some additional research and presents you with the following list of information about the process. Please rank each of these pieces of information to indicate its importance, with 1 being "essential," 2 being "secondary consideration," and 3 being "minor detail."

11. What changes, if any, would you make to the process in light of the most essential pieces of information your intern provided to you?

Corrected and more efficient block flow diagram is now provided.

12. After considering the available information, one of your colleagues suggests two improvements which are incorporated into the above design. Will you incorporate these suggestions? Why or why not?

13. After considering all available information, describe any and all changes that you would make to the original design the intern presented to you (shown above). Justify all of the changes that you made to your manager.

Figure 1. Outline of assessment design.^[48]

develop a rubric. All three experts had decades of industry experience, as well as several years of experience teaching capstone design courses at research universities. From these expert responses we developed a rubric to determine how expert-like students' solutions were (see Table 1).

ANALYSIS

For the quantitative analysis, we calculated students' assessment scores according to the scoring scheme summarized in Table 2. The assessment has eight quantitative measures of expert-like performance. The Evaluation Criteria, Information Request, and Safety scores represent the number of factors in Questions 1, 9, and 8 respectively that were mentioned by experts and students, divided by the total number of factors mentioned by students. These scores range from 0-100% because the total number of expert-like responses cannot exceed the total number of student respones. On Question 10, experts were unanimous in citing three pieces of information as essential to the process. The Information Ranking score is the fraction of those three pieces of information that students ranked as important. We do not consider the ordinal ranking of all information, just the agreement on the top 3 pieces of information as identified by experts. That is, if a student ranked two of the three top expert pieces of information as essential, they would receive an information ranking score of 67%. There were three errors in the original design shown to students. Design Errors is the fraction of those three errors that students noticed. Similarly, there were four improvements we identified that students could suggest. Design Improvements is the fraction of the four objective improvements that could be made to the corrected process that students correctly suggested. Students did not suggest any improvements outside of the ones we identified. Feasibility is the fraction of students who said the original design shown to them (with significant flaws) was feasible when asked in Question 3. Improved Process is the fraction of students who accepted the process improvements shown to them in Question 12. We note that the first six measures are student-level scores, whereas the final two measures are course-level scores. We use these different measures of expertise because the different sections of the assessment measure different types of problem-solving. A coded example solution may be found in the Appendix.

For Evaluation Criteria, Information Request, Information Ranking, Design Errors, and Design Improvements, we used a paired, two-tailed t-test for unequal variances to determine if these differences were statistically significant. We then calculated Cohen's d as a measure of the gains that students made from pre- to post-test. Cohen's d measures the difference between the means of two distributions in units of standard deviations of those distributions. Effect sizes 0.5 or greater are considered to be large, and effect sizes of 0.2 - 0.5 are

126

medium.^[56] We did not calculate Cohen's *d* for Safety scores because Question 8 was located at the end of the pre-test. We found that this question placement biased student responses to only mention the safety issues present in the information given to the students in Part 3 of the assessment, so the measures did not reflect students' own knowledge. On the binary measures of how many students thought the original design was feasible, and how many students accepted the improvements suggested to the design, we used a chi-squared test to determine if there were statistically significant differences in counts. A chi-squared test is appropriate for counts because it is used to measure the difference between expected frequencies of an event and the observed frequencies; in this case the event is whether a student indicated the process was feasible or whether they accepted the improved design.

RESULTS

We found statistically significant increases in Evaluation Criteria (15% pre, 25% post, t(89) = 2.89, p = 0.002, d = 0.35, medium effect size), Information Request (39% pre, 59% post, t(89) = 3.02, p = 0.002, d = 0.37, medium effect size), Information Ranking (51% pre, 59% post, t(89) = 1.76, p = 0.041, d = 0.25, medium effect size), Design Errors (8.1% pre, 18% post, t(89) = 2.99, p = 0.002, d = 0.37, medium effect size), and Design Improvements (12% pre, 27% post, t(91) = 5.50, p < 0.001, d = 0.65, large effect size) scores.

The change in Evaluation Criteria scores is due to students mentioning more criteria that were cited by experts, not citing fewer overall criteria. Compared with their pre-test responses, students were more likely to mention safety as an important criterion (13% pre, 29% post), as well as product yield (12% pre, 21% post), the quality of the separations processes (5.5% pre, 16% post), and the complexity of the process (8.8% pre, 15% post). Students were less concerned with issues related to the visual representation of the process (25% pre, 8.8% post). Students also became more aware of issues related to process economics (45% pre, 60% post), energy use (18% pre, 52% post), and material re-use (38% pre, 59% post) from pre- to post-test, though these were not criteria cited by the experts.

Analysis of students' responses to Question 10 reveals three notable changes in the information that students requested at pre-test versus post-test. At post-test, students again were much more interested in safety information, e.g. material toxicity (3.3% pre, 25% post) and quantitative information on process flow rates (5.5% pre, 22% post).

The difference in students' information ranking scores can be explained by a shift from students thinking about the three most essential pieces of information as "secondary considerations" to viewing them as "essential" information. Students were more likely to rank information related to environmental hazards (54% pre, 65% post) and the quality of separations

	TABLE 1
Request, and Sa	oring the assessment based on the responses of three experts. For Evaluation Criteria, Information afety, the rubric items were mentioned by at least one of the three experts. For Information Request, perts agreed that these items were essential pieces of information. The Design Errors and Design Improvements were predetermined based on the design of the assessment.
Topic Area	Elements of Expert Solution
Evaluation Criteria	 Complexity: whether the process accomplishes its objectives with a minimal number of units Safety: whether there are particular safety considerations necessary to accommodate this process Stream compositions: whether mole or mass fractions of components in each stream are listed Flowrates: whether the flowrates are specified to determine production volume and equipment sizing Yield: whether yield of the desired product is given Reasonable Separations: whether the specified separations follow the physical properties of the given chemicals and are feasible Mass balance: whether the overall mass balance is satisfied Metallurgy: whether consideration is given as to what materials to make units and pipes out of, given high temperature and corrosive environment Process failure: what particular failure modes the process has
Information Request	 Flow rates: requesting that the flow rates of all streams be given Equipment Specs.: requesting information on the sizing and construction of the unit operations Yields: requesting the yield of the final product Temperature/Pressure: requesting operating temperatures and pressures of streams and units Complexity: asking whether the same objectives could be accomplished with fewer units Safety: asking what the safety considerations on materials handling and process operation are Heat management: asking what is the plan for heating and cooling various parts of the process Waste/environment: asking what is the waste disposal plan Byproduct: asking about potential uses, recycles, or markets are for various by byproducts
Information Ranking	 Chlorine gas is extremely corrosive at high temperatures and can corrode as much as 1/2" steel per year Carbon tetrachloride is toxic and can cause extreme liver toxicity Carbon tetrachloride (reactant) can be reformed from chlorine gas and hexachloroethane (byproducts) in the presence of an antimony pentachloride catalyst
Design Errors	 Carbon tetrachloride is recycled into a separator instead of the initial reactor, creating an infinite accumulation loop The recycle stream for carbon tetrachloride lacks a purge, leading to the buildup of impurities Chlorine gas is missing from the effluent of the second reactor
Design Improvements	 Hexachloroethane can be recycled to create more product The separations can be reordered to more efficiently use heating and cooling The number of units can be reduced by combining redundant separation operations The large temperature swings can be used to heat other parts of the process
Safety	 Chlorine and carbon tetrachloride are toxic and bad for the environment and plant workers The large recycle streams could lead to buildup of impurities and process failure Chlorine could corrode pipes and units and cause leaks There could be runaway reactions which lead to combustion There should be a complete HAZOP analysis performed The process temperatures are extremely high

	TABLE 2	
Quantitative me	easures for scoring the assessment. All scores range from 0 to 100 %. The contribut refer to the questions listed in Figure 1.	ing questions
Topic Area	Description	Contributing Questions
Evaluation Criteria	Number of student-identified criteria that were also cited by experts divided by the total number of student-identified criteria.	1
Information Request	Number of student-requested pieces of information that were also requested by experts divided by the total number of student-requested pieces of information.	9
Information Ranking	Experts cited three pieces of information as most important. This score is the fraction of those pieces that students rank as most important.	10
Design Errors	Fraction of errors that students notice in the design.	1 - 4
Design Improvements	Fraction of improvements that students suggest to the design.	1 - 7
Safety	Number of student-identified safety considerations that were also cited by experts divided by the total number of student-identified safety considerations.	8
Feasibility	Fraction of students who say the original design is feasible.	3
Improvements	Fraction of students who accept the improvements suggested to the design.	12

(31% pre, 42% post) as essential at post-test and were less concerned with the cost of ancillary equipment (31% pre, 15% post). At post-test, students were more likely to suggest changes to the process based on the information they had to rank: 36% of pre-test responses suggested no changes to the process compared with 26% of post-test responses. However, their changes were largely superficial, either expressing generic concerns about process safety (6.6% pre, 18% post) and ways to avoid the use of a toxic reactant (7.7% pre, 26% post).

At post-test, students were far more likely to notice the missing purge stream, which would result in infinite mass accumulation in the process (1.1% pre, 32% post). However, they were not any better at noticing the other errors in the design. They were far more likely to suggest two design improvements: recycling unused intermediate (29% pre, 48% post) and adding units to recover lost heat (7.7% pre, 43% post). Students were less likely to suggest recycling unreactive byproducts or desired product at post-test (26% pre, 13% post), which would make the process worse. They were more likely to suggest changing process temperatures to save energy (4.4% pre, 14% post), which is well intended but has economic implications about the size of the equipment that students did not consider when suggesting this change.

The safety concerns that students had about the process at post-test were overwhelmingly about the hazards associated with using chlorine gas (44%) and the high temperatures in the process (34%) without explaining why they were concerned about those issues. A small fraction of students elaborated that they were concerned about chemical spills and disposal (6.6%) or reactor combustion (7.7%). About 14% of students

expressed general concerns about using hazardous chemicals but did not specify which chemicals or what their concerns were. We note that safety concerns frequently appeared in students' requests for additional information and criteria for evaluating the process, not only in response to the safety question.

There was no statistically significant change in the number of students who said the original design was feasible (59 % at pre-test, 64% at post-test, p = 0.50). Note that there are no standard errors on this measurement because it is a frequency. Students who said the process was feasible generally shifted their reasoning from "it produces the desired product" (43% pre, 26% post) to saying that the process would work but it would be inefficient (15 % pre, 36% post). Students who said the process was not feasible were more likely to cite the missing purge stream in the original design (2.8% pre, 34% post).

There was no statistically significant change in the number of students who accepted the proposed changes to the process at the end of the assessment (87 % pre, 80 % post, p = 0.18). Students who accepted the changes were more likely to cite the fact that the proposed design had fewer units (19% pre, 36% post), saved energy (7.7% pre, 21% post), and reduced cost (10% pre, 18% post). Notably, only 35% of the students who accepted the changes because there were fewer units also identified that this would lead to cost and energy savings. Students who did not accept the changes were more likely to say that the increased size of the units outweighed the cost savings from recycling intermediates (8.3% pre, 16% post) and that they objected to there being product in the recycle/ purge stream (8.3% pre, 26% post). Notably, 25% of pre-test responses who did not accept the changes said the process produced no product, while none of the students said this at post-test.

DISCUSSION

Broadly speaking, we see that students' problem-solving improves over the cornerstone design course. The improvements seem to be in areas where they are given practice in making certain expert decisions during the cornerstone design course, and that the problem-solving gains are smaller when this practice is more limited. For example, students get substantial practice deciding what information is needed to solve the problem and how important various pieces of information are during the final project. Students get some practice on the homework reflecting on how well a given solution holds, but they clearly need more practice in this area. Qualitative analysis of students' open-ended responses reveal that students' predictive frameworks are starting to more closely resemble experts', particularly when it comes to issues of process safety. Despite these gains, there are also areas where students lack clear connections between important features of the problem.

We also note that the problem-solving scores of first-year students at the end of their first engineering design course are similar to the problem-solving scores of senior students entering the capstone design course.^[51] This suggests that students may not be getting the chance to practice making many of the expert problem-solving decisions during the engineering science courses they take in the second and third years of the curriculum. Another possibility is that the problem-solving skills they do learn in these courses do not transfer to other contexts (e.g. chemical process design) because they are developing different predictive frameworks in those courses. There are also limitations on the comparison between these students and the seniors studied.^[51] The measurements are not longitudinal data, so we cannot say for sure that the students we studied here would receive similar scores on the assessment when they began their capstone design courses. Furthermore, the seniors in this study [51] were from a different university. However, the two universities are both highly selective research intensive universities with similar demographics, both have a cornerstone chemical process design course, and both have similar curricula.

Changes Reflecting Practice Making Decisions

Many of the improvements in problem-solving that we see may be explained by students getting practice making relevant expert decisions during the cornerstone design course. The improvement in Evaluation Criteria scores suggests that students are getting practice defining what are the goals of the problem, as well as what are important features of the problem. Students get targeted practice defining the problem goals when they complete the "redefine the problem" exercises.^[53] We also assert that redefining a problem requires students to identify the important features of that problem. Students only get a very limited opportunity to practice this at the beginning of the cornerstone course.

The skills practiced during the final design project seem like a plausible explanation for the increase in Information Request and Information Ranking scores. During the final project in particular, students had the opportunity to practice deciding what information was important and relevant to solving the problem. Though they were given a limited set of sources and thus had to do a little of their own research, the sources provided far more information than was necessary to solve the problem they were interested in. Thus, they had to parse the documents for what information was needed to solve the problem and prioritize what information would be most important for reaching a solution. The final project was relatively structured, so students did not have to decide what conclusions could be drawn based on this new information, only whether or not the information was needed and how important it was.

We note that two large improvements were on students' Design Improvements scores (15% change, d = 0.65) and Design Errors scores (10% change, d = 0.37). The first is a large effect size, and the second is medium. In terms of the decisions made, these scores primarily measure how well students are deciding how well a given solution (the flawed design) holds. Students receive scaffolded practice in making these decisions when they complete the process analysis exercises from Ref.^[55], which suggests that practicing these kinds of troubleshooting exercises better enables students to reflect on a given solution. Because of the limitations of the assessment, we cannot determine whether this reflective practice would transfer to a student analyzing their own designs, but this seems plausible. Future investigations should determine whether troubleshooting practice allows students to better reflect on their own solutions. Finally, we note that, despite the students' improvements on Design Improvements and Design Errors scores, there is still room for growth. The majority of students were unable to identify the mass accumulation problem and thus predicted that the original flawed design was feasible. Based on previous work in mechanical engineering design, we hypothesize that deliberate practice with reflecting on their problem-solving processes and solutions might help students improve further.^[57]

Developing Predictive Frameworks

The analysis of students' open-ended solutions reveals how students' predictive frameworks are developing over the course of the cornerstone design experience. For example, students were less likely to confuse chemical species at posttest (reflected by suggesting that waste products or desired products be recycled). We also see students progressing towards expertise by becoming more quantitatively oriented in their evaluation of the process and the information they request. Flow rates, operating parameters, and product yields are all quantitative measures that help experts judge the feasibility and quality of a chemical process. As a result of practice using such measures to evaluate processes in their homework exercises, this area of students' predictive frameworks becomes more developed during the course.

One theme where the students see consistent shifts toward expertise is the role of safety in chemical process design. Students were more likely to cite safety as an important criterion for evaluating the initial process, more likely to request safety information, and more likely to rank safety information as essential. Safety was a theme that was consistently emphasized during instruction. For example, issues of mass accumulation and byproduct stream contamination were often framed in terms of safety hazards - e.g. what is an acceptable level of pollutant in a waste stream from a process that the students are designing. Though there were significant shifts toward thinking about safety, this area of students' predictive frameworks is still being developed. For example, students' concerns about process safety were largely about the high process temperatures and that chlorine is hazardous. While both of these things are true, and both were mentioned by experts, very few students articulated the consequences of these facts - namely that the plant is at increased risk of combustion/runaway reaction because of these high temperatures, and the plant workers are at increased risk of hazardous chemical exposure.

There are other areas where development of predictive frameworks is evident. In Question 12 we see far more students citing the reduced number of units as a reason that the improvements to the process should be accepted. However, only 34% of students who identify the reduced number of units as an improvement to the process further articulate that this results in energy and cost savings. This reflects an area where students' predictive frameworks may lack strong relationships between important features of this problem. As a further example, the students who reject the improvements on the basis that the size of the units is greatly increased are able to make predictions about how adding recycle streams will affect the process, but they have a weak understanding of the relationship between the size of the unit and its cost (which is usually sublinear).

It is worth noting that the development of students' predictive frameworks is not a guaranteed outcome of instruction. Indeed, in a separate study we found that few students who had taken an introductory physics course had developed predictive frameworks that were sophisticated enough to allow them to plan their problem-solving strategy in advance.^[57] Even in a course where instruction was carefully designed to teach problem-solving, that study showed that we failed to help students develop more expert-like predictive frameworks.^[58]

Limitations and Future Work

We cannot prove that the structure and content of the course are the cause of the shifts in problem-solving we see here. However, based on research that describes the development of expertise,^[54] it seems plausible that the course is indeed helping students developing these problem-solving skills through deliberate practice. Deliberate practice is carefully scaffolded, targeted practice of component skills needed for expert performance.^[59] Through this practice and feedback from an expert teacher or coach, the student develops the tools they need for consistent expert performance in a given domain. Future studies should administer this same problemsolving assessment as a pre- and post-test in a typical mass and energy balances course that does not have a significant design component. That would provide a pseudo-control group that would help show that the gains we see here are indeed the result of instruction in this course.

Another limitation is that we only investigated problemsolving in a single cornerstone design course in a single discipline. Future work should focus on developing disciplinespecific assessments so that problem-solving can be measured in other contexts. This would allow us to determine whether cornerstone design courses, in general, teach students how to solve problems, or whether this is a unique instance.

This study suggests that students' problem-solving in the relatively broad areas probed by this assessment is largely stagnant between the end of the cornerstone design experience and the beginning of the capstone design experience. While many core engineering courses provide extensive coverage of chemical engineering content, they often offer very little practice in making design decisions using this content. We are currently working with instructors of engineering science courses in years 2 and 3 of the curriculum to develop interventions that improve students' problem-solving during the middle years.

CONCLUSIONS

We administered an assessment of problem-solving as a pre- and post-test in a cornerstone chemical engineering design course. We find that students' problem-solving skills improved over the course of a cornerstone design experience in chemical engineering. Students saw particularly large gains suggesting improvements to the block flow diagram and identifying what information is necessary and important to solve the problem. This can be plausibly explained by the practice students received making these decisions during the final project and homework assignments in the course. Students received less practice identifying important features of the problem and are thus less expert-like in this area at the end of the course.

We also see evidence that the cornerstone design course helps students develop more expert-like predictive frameworks. Students are better at keeping track of important features of the problem and become particularly well aligned with experts on issues of process safety. Despite this progress, there are also areas where students' predictive frameworks are still developing, as evidenced by weak links between important features of the problem at hand. This progress toward expertise is notable, as developing expert predictive frameworks is not a guaranteed outcome of introductory STEM courses, even when they are designed to teach problem-solving.^[58]

REFERENCES

- ABET, Engineering Accreditation Commission (2019) Criteria for Accrediting Engineering Programs 2019-2020. <u>https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2019-2020/</u> accessed March 4, 2021.
- Dym CL (2004) Design, systems, and engineering education. International Journal of Engineering Education. 20(3): 305–12. <u>https://</u> www.ijee.ie/articles/Vol20-3/IJEE2517.pdf
- Dym CL, Agogino AM, Eris O, Frey DD, and Leifer LJ (2005) Engineering design thinking, teaching, and learning. *Journal of Engineering Education* 94 (1): 103–20.
- Jonassen DH, Strobel J, and Lee CB (2006) Everyday Problem Solving in Engineering: Lessons for Engineering Educators. *Journal of Engineering Education* 95 (2): 1–14.
- Jonassen DH (2012) Design for decision making. Educational Technology: Research and Development 60: 341–59. <u>https://doi.org/10.1007/s11423-011-9230-5</u>
- Dutson AJ, Todd RH, Magleby SP, and Sorensen CD (1997) A review of literature on teaching engineering design through project-oriented capstone courses. *Journal of Engineering Education* 86 (1): 17–28. https://doi.org/10.1002/j.2168-9830.1997.tb00260.x
- Dym CL (1999) Learning engineering: Design, languages, and experiences. Journal of Engineering Education 88 (2): 145–48. <u>https://doi.org/10.1002/j.2168-9830.1999.tb00425.x</u>
- Sheppard S, and Jennison R (1997) Freshman engineering design experiences and organizational framework. *International Journal of Engineering Education* 13 (3): 190–97.
- Ford LP, Brennan J, Cole J, Dahm KD, Jamieson MV, Landherr LJ, Silverstein DL, Vaughen BK, Vigeant MA, and Thiel SW (2020) How we teach: Chemical engineering in the first year. *Proceedings of the ASEE Annual Conference*, available at <u>https://www.asee.org/public/ conferences/172/papers/30199/view</u>
- Allam YS, Whitfield CA, and Phanthanousy JN (2012) Scaffolding provided to engineering students in cornerstone design project scenarios related to practices of expert designers. *Proceedings of* the 2012 ASEE Annual Conference and Proceedings. <u>https://doi.org/10.18260/1-2--21898</u>
- Bodnar LJ, Lagoudas MZ, Hodge JQ, Smith TA, Orozco JA, Corso JG, and Cortes I (2012) Engaging freshman in team-based engineering projects. *Proceedings of the 2012 ASEE Annual Conference and Proceedings*. <u>https://doi.org/10.18260/1-2--21288</u>
- Cole RB, Gallois B, and Sheppard K (1999) Redesigning the first year engineering curriculum. *Proceedings of the 1999 Annual Conference*, *ASEE*, available at <u>https://peer.asee.org/7919</u>
- 13. Pavelich MJ, Olds BM, and Miller RL (1995) Real-world problem solving in freshmen-sophomore engineering. *New Directions*

for Teaching and Learning (61): 45–54. <u>https://doi.org/10.1002/</u> tl.37219956108

- Petroski H (1999) Polishing the gem: A first-year design project. Journal of Engineering Education 87 (4): 445–49. <u>https://doi.org/10.1002/J.2168-9830.1998.TB00377.X</u>
- Safferman SI, Zoghi M, and Farhey DN (2001) First year civil and environmental engineering design experience. *Journal of Engineering Education* 90(4): 645–51. <u>https://doi.org/10.1002/j.2168-9830.2001</u>, <u>tb00654.x</u>
- Swenson J, Potsmore M, and Danahy E (2014) Examining the engineering design process of first-year students during a hands-on, in-class design challenge. *Proceedings of the 2014 ASEE Annual Conference & Exposition*. <u>https://doi.org/10.18260/1-2--20449</u>
- Atman CJ, Chimka JR, Bursic KM, and Nachtmann HL (1999) A Comparison of freshman and senior engineering design processes. *Design Studies* 20 (2): 131–52. <u>https://doi.org/10.1016/S0142-694X(98)00031-3</u>
- Dally JW, and Zhang GM (1993) A freshman engineering design course. *Journal of Engineering Education* 82(2): 83–91. <u>https://doi.org/10.1002/j.2168-9830.1993.tb00081.x</u>
- Dym CL (1994) Teaching design to freshmen: Style and content. Journal of Engineering 83(4): 303-310. <u>https://doi.org/10.1002/j.2168-9830.1994.tb00123.x</u>
- Booth J, Jon M, and Doyle TE (2012) Importance of first-year engineering design projects to self-efficacy: Do first-year students feel like engineers? *Proceedings of the Canadian Engineering Education Association*. https://doi.org/10.24908/pceea.v0i0.4650
- Marra RM, Palmer B, and Litzinger TA (2000) The effects of a firstyear engineering design course on student intellectual development as measured by the Perry Scheme. *Journal of Engineering Education* 89 (1): 39–45. https://doi.org/10.1002/j.2168-9830.2000.tb00492.x
- Agogino AM, Sheppard S, and Oladipupo A (1992) Making connections to engineering during the first two years. Proceedings of the Twenty-Second Annual conference Frontiers in Education. <u>https://doi.org/10.1109/FIE.1992.683462</u>
- Knight DJ, Carlson LE, and Sullivan JF (2007) Improving engineering student retention through hands-on, team-based, first-year design projects. Proceedings of the International Conference on Research in Engineering Education.
- Shin NS, Jonassen DH, and McGee S (2003) Predictors of wellstructured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science Teaching* 40 (1): 6–33. <u>https:// doi.org/10.1002/tea.10058</u>
- Jonassen DH (2014) Ch 6: Engineers as problem solvers. Johri A and Olds BM (Eds). *Cambridge Handbook of Engineering Education Research*. Cambridge University Press. New York, NY/USA.
- Means B, Salas E, Crandall B, and Jacobs TO (1993) Training decision makers for the real world. Klein GA, Orasanu J, Calderwood R, and Zsambok CE (Eds). *Decision Making in Action: Models and Methods*. Ablex Publishing. New York City, NY/USA.
- Woods D (2000) An evidence-based strategy for problem solving. Journal of Engineering Education 89(4): 443–59. <u>https://doi.org/10.1002/j.2168-9830.2000.tb00551.x</u>
- Chi MTH, Feltovich PJ, and Glaser R (1981) Categorization and Representation of physics problems by experts and novices. *Cognitive Science* 5(2): 121-152. <u>https://doi.org/10.1207/s15516709cog0502_2</u>
- Duncker K, and Lees LS (1945) On Problem-Solving. Psychological Monographs 58(5): i-113. <u>https://doi.org/10.1037/h0093599</u>
- Newell A, and Simon HA (1972) Human Problem Solving. Prentice-Hall. Upper Saddle River, NJ/USA.
- Adams WK and Wieman CE (2015) Analyzing the many skills involved in solving complex physics problems. *American Journal of Physics* 83: 459. <u>https://doi.org/10.1119/1.4913923</u>
- Heller P and Hollabaugh M (1992) Teaching problem solving through cooperative grouping part 2: Designing problems and structuring groups. *American Journal of Physics* 60 (7): 637–44. <u>https://doi.org/10.1119/1.17118</u>

- Heller P, Keith R, and Anderson S (1992) Teaching problem solving through cooperative grouping part 1: Group versus individual problem solving. *American Journal of Physics* 60: 627. <u>https://doi. org/10.1119/1.17117</u>
- Heller JI and Reif F (1984) Prescribing effective human problemsolving processes: Problem description in physics. *Cognition and Instruction* 1(2): 177-216. <u>https://doi.org/10.1207/s1532690xci0102_2</u>
- Hsu L, Brewe E, Foster TM, and Harper KA (2004) Resource letter RPS-1: Research in problem solving. *American Journal of Physics* 72: 1147. https://doi.org/10.1119/1.1763175
- Schoenfeld HA (1985) Mathematical Problem Solving. Elsevier. Amsterdam, Netherlands.
- 37. Polya G (2004) *How to Solve It: A New Aspect of Mathematical Method.* Princeton University Press. Princeton, NJ/USA.
- Reif F, and Heller JI (1982) Knowledge structures and problem solving in physics. *Educational Psychology* 17(2): 102–27. <u>https://doi.org/10.1080/00461528209529248</u>
- Miller GA (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63(2): 81–97. <u>https://doi.org/10.1037/h0043158</u>
- Swelle, J (1988) Cognitive load during problem-solving: Effects on learning. *Cognitive Science* 12(2): 257–85. <u>https://doi.org/10.1207/ s15516709cog1202_4</u>
- Dhillon AS (1998) Individual differences within problemsolving strategies used in physics. *Science Education* 82(3): 379–405. <u>https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-</u> 237X(199806)82:3%3C379::AID-SCE5%3E3.0.CO:2-9
- Larkin J, McDermott J, Simon DP, and Simon HA (1980) Expert and novice performance in solving physics problems. *Science* 208 (4450): 1335–42. https://doi.org/10.1126/science.208.4450.1335
- Simon DP and Simon HA (1978) Individual differences in solving physics problems. *Children's Thinking: What Develops?* RS Siegler (Ed). Lawrence Erlbaum Associates. Mahwah, NJ/USA.
- Singh C (2002) When physical intuition fails. American Journal of Physics 70: 1103–9. <u>https://doi.org/10.1119/1.1512659</u>
- Chi MTH, Bassok MW, Lewis MW, Reimann P, and Glaser (1989) Self-Explanations: How students study and use examples in learning to solve problems. *Cognitive Science* 13(2): 145–82. <u>https://doi.org/10.1207/s15516709cog1302_1</u>
- Dufresne RJ, Gerace WJ, Hardiman PT, and Mestre JP (1992) Constraining novices to perform expertlike problem analyses: Effects on schema acquisition. *Journal of the Learning Sciences* 2(3): 307–31. https://psycnet.apa.org/doi/10.1207/s15327809jls0203_3
- 47. Čančula MP, Planišič G, and Etkina E (2015) Analyzing patterns in experts' approaches to solving experimental problems. *American*

Journal of Physics 83(4): 366-74. https://doi.org/10.1119/1.4913528

- 48. Price AM, Kim C, Burkholder EW, Fritz AV, and Wieman CE (2021, in press) A detailed characterization of the expert problem-solving process in science and engineering: Guidance for teaching and assessment.
- Crandall B, Klein GA, and Hoffman R (2006) Working Minds: A Practitioner's Guide to Cognitive Task Analysis. MIT Press. Cambridge, MA/USA.
- Burkholder EW, Price AM, Flynn MP, and Wieman CE (2019) Assessing problem solving in science and engineering programs. 2019 Perc Proceedings. <u>https://doi.org/10.1119/perc.2019.pr.Burkholder</u>
- Burkholder EW and Wieman CE (2020) Comparing problemsolving across two capstone design courses. *Proceedings of the 2020 IEEE Frontiers in Education Conference*. <u>https://doi.org/10.1109/</u> FIE44824.2020.9273820
- Burkholder EW and Wieman (2020) Comparing problem-solving across capstone design courses in chemical engineering. *Proceedings* of the 2020 IEEE Frontiers in Education. <u>https://doi.org/10.1109/ FIE44824.2020.9273820</u>
- 53. Burkholder EW and Wieman CE (2020) Testing an assessment of problem solving in introductory chemical process design courses. 2020 ASEE Annual Conference Proceedings.
- 54. Ericsson KA (2006) The influence of experience and deliberate practice on the development of superior expert performance. Ericsson KA, Charness N, Feltovich PJ, and Hoffman RR (Eds). *Cambridge Handbook of Expertise and Expert performance*. Cambridge University Press. New York City, NY/USA.
- Duncan TM. and Reimer JA (2019) Chemical Engineering Design and Analysis: An Introduction. Cambridge University Press. New York City, NY/USA.
- 56. Fraenkel JR, Wallen NE, Hyun HH (2010) *How to Design and Evaluate Research in Education*. McGraw-Hill. New York, NY/USA.
- Salehi S (2018) Improving Problem-Solving Through Reflection. PhD thesis, Stanford University. <u>https://search.proquest.com/open_view/5daf498df36cb446000d4372e02c4bcc/1.pdf?pq-origsite=gsch_olar&cbl=18750&diss=y</u>
- Burkholder EW, Miles JK, Layden TJ, Wang KD, Fritz AV, and Wieman CE (2020) Template for teaching and assessment of problemsolving in introductory physics. *Physical Review Physics Education Research* 16(1): 010123. <u>https://doi.org/10.1103/PhysRevPhysEducRes.16.010123</u>
- 59. Ericsson KA, Krampe RT, and Tesch-Romer C (1993) The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 100(3): 363–406. <u>https://psycnet.apa.org/ doi/10.1037/0033-295X.100.3.363</u> □

APPENDIX : Coded Example Solution

Question Number	Student Response	Codes	Score
1	1. Efficiency - how much product it pro-	Expert Criteria: Process	Evaluation Crite-
	duces based on input, how much energy	safety, Product yield. Stu-	ria Score: 0.40
	it requires 2. Return on investment - is	dent Criteria: Process	
	the process economical? 3. Safety - how	economics, energy used,	
	safe is this process, how much waste is	minimize waste	
	produced 4. Environmental - utilize heat		
	exchangers and recycling streams. I be-		
	lieve the most important part of a design		
	is efficiency and its viability, in which the		
	design must be able to produce the de-		
	sired product at a decent rate compared		
	to laboratory batch processes. From then		
	on, it must be an economical process or		
	it will not be sustainable in the long run.		
	Other factors that must be included are		
	how safe it is and how optimized it is to		
	prevent harm to human health and envi-		
	ronmental damage.		
2	No. CCl4 is subject to accumulation. The	Design Errors: CCl4 purge	Design Errors
	liquid mixture exiting the liquid gas sepa-	missing. Design Changes:	Score: 0.33
	rator at 100C can immediately be placed	Remove necessary separa-	
	in a pyrolysis reactor to increase separa-	tor.	
	tion efficiency.		
3	Yes, but it is redundant and inefficient.	Reasons Feasible: Ineffi-	Feasible (1)
		cient	
4	[Not shown to student]		
5	The pyrolysis reactor II is still inefficiently	No new codes due to vague	
-	placed. More recycle streams of CCI3-	response.	
	CCI3 can be added.	responser	
6	No, it is missing recycling and heat ex-	No new codes due to vague	
•	changers.	response.	
7	I would move the pyrolysis reactor II	Design Improvements:	Design Improve-
/	to immediately after the 100C liquid	CCI3-CCI3 recycle, Add	ments Score:
	gas separator. The output can then be	heat exchangers.	0.50
		fieat excitatigers.	0.00
	separated, with CCI3-CCI3 recycled and		
	purged into the pyrolysis reactor II. Heat		
	exchangers could also be added where		
	the output from pyrolysis reactor l is heat		
	exchanged to the output from the 100C		
0	liquid gas separator.	Sefet a CIO harrowith an	Cafatu C
8	Chlorine gas is corrosive and a health haz-	Safety: Cl2 hazards, corro-	Safety Score:
	ard to both the humans and the environ-	sion	0.50
0	ment.	Forest late 2 P	1-6 D :
9	Energy consumption, size of streams and	Expert Information Re-	Info. Request
	reactors to consider economic return on	quested: Energy consump-	Score: 0.75
	investment. This would allow us to make	tion, flow rates, equipment	
	better conservation of energy and calcu-	specifications. Student	
	late the economic viability of such a reac-	Information Requested:	
10	tor design.	Process economics.	1
10	Ranking task, no ambiguity in coded re-		Information
	sponses.		Ranking Score:
4.4		<u> </u>	0.67
11	Ensure the safety of the design.	Safety - generic.	
12	Yes, it makes the process more efficient.	Reasons to Accept: Effi-	Accept (1)
	However, the economic viability of imple-	ciency. Reasons to Reject:	
	menting more recycle streams must be	Increased unit sizes.	
	considered as reactor and piping must		
	be larger to accommodate the increased		