ASME Journal of Mechanical Design
Online journal at:
https://asmedigitalcollection.asme.org/mechanicaldesign

A Comparative Analysis of Cognitive Engagement Through Embedding Design-Based Learning in Fluid Mechanics

Boni Yraguen[1]
Department of Mechanical Engineering,
Pennsylvania State University,
State College, PA 16802
e-mails: bjy5264@psu.edu;
boni.yraguen@vanderbilt.edu

Elisa Koolman
Department of Mechanical Engineering,
University of Texas at Austin,
Austin, TX 78712
e-mail: elisa.koolman@my.utexas.edu

Anna Lummus
Department of Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332
e-mail: alummus6@gatech.edu

Katherine Fu
Department of Mechanical Engineering,
University of Wisconsin—Madison,
Madison, WI 53706
e-mail: kfu26@wisc.edu

Roxanne Moore
College of Lifetime Learning,
Georgia Institute of Technology,
Atlanta, GA 30332
e-mail: roxanne.moore@gatech.edu

Fluid mechanics is an early mechanical engineering course where abstract concepts gain physical, applicable meaning. It is therefore a prime venue to teach higher-order engineering skills—problem definition, modeling, and solution analysis—needed for the ill-structured, open-ended problems engineers face in practice. These skills are primarily taught in design courses which account for only 20% of the curriculum. Thus, design thinking and pedagogy are needed in core engineering courses taught throughout the degree. This study evaluates the impact of an authentic learning assignment, titled design your own problem (DYOP), on students' higher-order thinking in a fluid mechanics course. Using Bloom's Taxonomy to gauge cognitive engagement, we conducted content analysis of students' cognitive reflections on typical engineering assessments (quizzes) and the DYOP. Results show a significant increase in higher-order cognitive skills during completion of DYOP compared with quizzes. This heightened engagement features greater analysis, evaluation, and creativity, indicating a shift toward more sophisticated problem-solving and metacognitive awareness among students. This pattern was consistent across diverse student cohorts, irrespective of gender, racial, or ethnic background, prior internship experience, or initial performance on quizzes. These findings present a simple, scalable, and effective method for incorporating higher-order cognitive skills into core courses within the engineering curriculum, thereby providing additional avenues of design-type training prior to students enrolling in design-focused courses such as senior design or capstone. Embedding authentic, open-ended tasks alongside traditional problems cultivates the modeling fluency, judgment, and analysis essential for engineering practice. [DOI: 10.1115/1.4070860]

Keywords: *design courses and curricula, cognitive-based design, design education*

Introduction

The widening gap between engineering curriculum and practice [1,2] has led to recent graduates facing challenges in navigating less-defined problem spaces, as noted by employers [3–5]. These graduates often require additional workplace training to acquire missing competencies [6]. Universities and engineering faculty members should be prepared to support these students in bridging that gap, as they do indeed possess the intelligence, determination, and other qualities needed to succeed [7]. One key issue driving this gap is tensions between convergent and divergent thinking in engineering training [8]. Many engineering curricula tend to emphasize "right answer" or convergent thinking [9]. This type of thinking is enforced by assessments that typically focus on problems with only one correct answer or solution. This emphasis on a single right answer may discourage creativity and unique problem-solving approaches, which could in turn limit the breadth of perspectives represented in engineering.

Within engineering curricula, design education is one of the only places where creative and divergent thinking is emphasized [10]. Divergent thinking is a type of open-ended problem-solving that involves producing a wide range of responses [11]. Most often, this type of thinking is taught through incorporating problem-based learning and open-ended questions. In an attempt to mitigate this issue, some engineering educators have turned to design-based education to "fix what is perceived to have been broken for many years: a content-focused curriculum, lacking in the context of real problems that would motivate and engage students and deepen their learning" [12]. Yet, many faculty teaching core engineering courses have limited formal training in design processes, methods, and tools [12]. Furthermore, project-based and design-focused learning initiatives [13–16] account for less than 20% of class time in engineering education [17]. Faculty from wide-ranging disciplines that teach core

[1]Corresponding author.

Contributed by the Design Education Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received August 20, 2025; final manuscript received December 29, 2025; published online March 9, 2026. Assoc. Editor: Nicholas Meisel.

engineering courses need an accessible pathway to incorporate divergent thinking and open-ended problem-solving into their courses.

Authentic learning assignments [18,19] may offer a practical way to begin addressing this gap in lecture-based courses. Authentic learning assignments are tasks that connect classroom learning to real-world situations, requiring students to apply their knowledge and skills to complex, meaningful problems that mirror those they might encounter outside of school. The goal of this study is to assess the effectiveness of an authentic learning assignment deployed in a core mechanical engineering class: fluid mechanics. The effectiveness of the assignment, titled design your own problem (DYOP), is assessed using content analysis to measure levels of learning as characterized using Bloom's Taxonomy [20,21]. The DYOP was designed to break the dichotomy between convergent and divergent thinking by encouraging creative and divergent thinking while reinforcing technical concepts that require convergent thinking. We hypothesize that the DYOP increases a student's divergent thinking (engagement with the subject matter material by means of upper levels of Bloom's Taxonomy: analyze, evaluate, and create) compared to that of a traditional quiz assessment. We evaluated this hypothesis using coded reflections capturing the student cognition of 54 students in a junior-level fluid mechanics class.

Background

Higher-Order Cognitive Skills in the Curriculum. The skills encompassed within the broader scope of critical thinking are inextricably tied to the design process. To ensure original and successful solutions, the design process, frequently characterized by problem definition, ideation, prototyping, and evaluation, depends on successfully applying critical thinking skills [22,23]. The engineering design process involves defining complex problems, exploring multiple solutions, and employing logical reasoning and abstraction techniques to address open-ended challenges [24]. It is understood that the ability to analyze and evaluate underscores the ability to problem-solve, make inferences, and draw conclusions based on evidence [25–27].

These design-based skills ground an engineer's ability to create innovative designs, yet creativity is underemphasized in engineering curricula [28–31]. Creativity can be broadly defined to encompass innovations or ideas that alter/grow a discipline and/or to simply include personally meaningful discoveries that occur while a student is learning [32]. As such, concepts of creativity could easily be scaffolded into the engineering curriculum.

These design and innovation skills can broadly be categorized into high-order cognitive skills like: analyze, evaluate, and create [21]. For example, a student's ability to correctly make assumptions and approximations falls within the category of analyze. Comfort with these analyze-level skills has been found to influence a student's performance in engineering design [33]. Furthermore, Girgis correlated these higher-order skills with the Accredidation Board for Engineering and Technology (ABET) criteria, emphasizing their importance in applying engineering knowledge postgraduation and fostering lifelong learning [34]. This correlation suggests that students engaging in higher levels of learning can more effectively apply engineering principles in their careers, with design and project-based courses naturally encompassing analyze, evaluate, and create levels. These higher levels are associated with critical thinking and design-based skills, yet typical engineering assessments found in most core courses (quizzes and exams) are limited in their ability to assess a student's ability in these areas [35].

The majority of scholarly articles focused on teaching these higher-order critical thinking and design-based skills in engineering education are based on studies implemented in or related to design-based courses [35]. Yet, in many engineering curricula, design is not explicitly taught until a student's final year (e.g., capstone or senior design courses). This may be too late to have an impact on student divergent thinking. Indeed, Leonard et al. found that capstone courses had little-to-no impact on student beliefs about creative and varied problem-solving strategies [36], while others have called for a scaffolded approach to teaching divergent thinking skills throughout the curriculum [10]. Some institutions, recognizing the importance of practical experiences to develop the necessary skills for effective thinking and problem-solving, emphasize experiential learning through internships and co-op experiences to introduce these skills earlier. As such, some curricula rely on co-op or internship programs to bridge the curriculum-practice gap [37,38] and hopefully expose students to these higher-order skills within real engineering contexts. Unfortunately, this strategy may not be feasible for schools with limited industry partnerships or for students due to location and timing constraints. This disproportionately affects rural and disadvantaged communities.

The Dichotomy Between Convergent–Divergent Thinking and Training in Engineering. With a significant amount of the engineering curriculum focused on providing technical content knowledge to students, it is natural that the more skill-based engineering training (e.g., problem identification and scoping, decision making, determining optimal solutions, etc.) gets siloed into design and experiential training. Furthermore, there are significant tensions between academic rigor and concepts like creativity in traditional engineering education which further cement this divide [8]. Traditional engineering instruction emphasizes academic rigor which results in an over-emphasis on convergent or "right answer" thinking [8]. One reason for this is the relative simplicity and ease of assessment and evaluation for closed-ended problems [11,39]. This type of problem is typically limited to assessing a student's capabilities only at the lower end of Bloom's Taxonomy (remember, understand, and apply) and, in some cases, can discourage students from attempting to creatively solve the problem [40]. Unfortunately, this emphasis on convergent thinking can contribute to students relying on familiar solution patterns that are sometimes removed from real engineering contexts [28,41].

On the contrary, open-ended questions encourage and emphasize creative and original thoughts from the respondents in their solutions and can measure the extent of the respondents' knowledge on a particular topic. In addition to that, open-ended problems concentrate on the process students utilize to solve problems, allowing students to cultivate creative thinking capabilities that will be useful in future courses or situations. Advantages include students becoming more actively engaged in learning and having numerous ways to solve problems, students expressing their opinions more freely through their solution strategies, and students developing original ideas and attaining higher critical and cognitive communication. These types of problems provide teachers with the opportunity to pose diverse questions to evaluate the student's depth of knowledge on a specific topic, while improving testing environments by reducing the propensity to cheat during an examination [40].

While the foundations of engineering knowledge are grounded in objective theory/physics, the practice of engineering and engineering design is inherently subjective with numerous valid pathways to problem-solve, analyze, and create engineering solutions and designs. Engineering curricula broadly reflect this dichotomy, with certain components focusing on the former and others emphasizing the latter. Instructional approaches that integrate both components are largely absent. There is a need in engineering education to break the dichotomy between convergent and divergent thinking in engineering problem-solving—teaching students when and how to differentiate and employ both types of thinking/problem-solving in their solutions [23].

Typical Core-Curriculum Interventions. Engineering education research is often focused on first-year or lower-level courses in the engineering track [18,34,42]. Upper-level engineering courses, such as fluid mechanics, build on foundational knowledge. They require students to apply this knowledge to more abstract concepts, which can then be used in a variety of physical applications. Fluid mechanics is typically one of the first courses mechanical engineering students encounter that requires this skill (e.g., it is one of the first courses where students must apply vector calculus, differential equations, boundary conditions, and force and mass balances simultaneously as appropriate based on real-world context). Thus, it is an ideal candidate for a lecture-based course in which alternative educational tools and interventions can be used to improve student engagement and learning at higher cognitive levels.

There have been prior efforts to study learning interventions in fluid mechanics [43–49]. These studies largely focus on problem-based learning [44,48], multidisciplinary lab experiments [45,47,49], and the use of computation as a teaching tool [46]. Prior interventions modify the course delivery style by flipping classrooms, incorporating many activities into lecture, designing complex projects for students to complete, and more. While these evidence-based practices typically have positive outcomes for student learning, the time and effort required by faculty to effectively implement them often present a significant barrier to adoption. Furthermore, most of these interventions still rely on closed-ended problems—where students execute a predetermined experiment or analysis that comes with a "correct" answer.

Theoretical Framework and Implications for the Study. Bloom's Taxonomy and the literature on reflections served as conceptual frameworks for this study. Here, we reference Bloom's *Cognitive* Taxonomy (as opposed to Bloom's Affective and Psychomotor Taxonomy) and will refer to it simply as "Bloom's Taxonomy" throughout this work. Bloom's Taxonomy was created to be a tool for measuring educational goals and standards [21]. There are six levels of cognition defined in order of increasing complexity: remember, understand, apply, analyze, evaluate, and create. This taxonomy is referenced in numerous educational studies as a common method for evaluating educational material and the level of cognitive engagement it enables [50–52]. Bloom's Taxonomy was selected as a conceptual framework for its simplicity, the familiarity that many engineering faculty have with it, and the ease with which engineering (in this case, fluid mechanics) learning objectives can be categorized into its various levels.

While Bloom's Taxonomy may provide a sufficient conceptual framework for categorizing student cognition, accurately capturing student engagement with the corresponding activities is complex. Indeed, students are likely unaware of the various levels of cognition that they engage in while working through a problem. Furthermore, typical engineering assessments result in completed exam solutions or project submissions—leaving out much of the thought and development process that took place to develop a "final product." Thus, this study relies on reflections to capture this process-based cognition. Within engineering education, reflection has emerged as a longstanding tool for enhancing student learning [53,54]. Even so, its application in technical engineering courses remains limited as educators employing reflections often focus on behavior-based aspects (like discouraging procrastination and encouraging thoughtful engagement with homework problem sets) rather than emphasizing traditional engineering content assessment [55,56]. Outside of engineering education, reflections are more often used to prompt higher-order thinking skills [54,57]. As such, multiple studies have offered categorical ways in which reflections can be coded [57–61]—highlighting their utility in capturing various forms of cognition that a person engages with during some process. Typically, these coding schemes focus on the student's ability to reflect (as reflection is commonly considered an advanced skill). While these categories may correlate to different levels of learning, they do not precisely evaluate the level or type of learning a student is engaging in with respect to the subject matter material. Other authors [58,59] have designed coding schemes that focus on the level of learning, as defined by Bloom's Taxonomy [20,21], to measure a student's cognitive process rather than their "depth" of reflection. Therefore, in this work, we employ reflections as a tool to capture the student's cognitive process during both quiz and DYOP assessments, then categorize that cognition into lower and higher levels of learning via Bloom's Taxonomy.

Pedagogical Design

This study was conducted in a fluid mechanics course in the mechanical engineering department at an R1 institution in the southeast United States in the Spring semester of 2021. The course was offered in a virtual format due to the COVID-19 pandemic. The format of delivery used a mixture of synchronous and asynchronous lectures. Theory and technical information were delivered using the asynchronous lectures in 5–15-min videos, and the synchronous course time was used primarily to complete example problems and facilitate discussion. The course included eight assessments: six quizzes and two projects. The projects were submitted as a midterm and a final. After each assessment, the students submitted a one-page reflection on their work within 24 h of completing the assignment. The reflection component of the assessment was graded for completion-only. The assessments and reflections were a mandatory component of the course regardless of a student's choice to participate in the current study. The format of the quizzes, the DYOP projects, and the reflection assignments will be described in detail in the following section.

The assessments throughout the course were broken up into two sections (i.e., two halves of one course). The first section included quiz 1 (Q1), quiz 2 (Q2), quiz 3 (Q3), and the midterm project. The second section included quiz 4 (Q4), quiz 5 (Q5), quiz 6 (Q6), and the final project. Each of the six quizzes corresponded to a concept or group of concepts in fluid mechanics: (1) hydrostatics, (2) Bernoulli's principle, (3) fluid kinematics, Reynolds transport theorem, and control volume analysis, (4) differential analysis of fluid flow, potential flow, and Navier stokes, (5) dimensional analysis and pipe flow, and (6) boundary layers and external flows. Students completed DYOP projects on the quiz topic for which the student received their lowest quiz grade. For example, if a student received 78% on quiz one (Q1), 95% on Q2, and 82% on Q3, the student would complete their midterm project on quiz topic one: hydrostatics. We then define Q1 (in this example) as the student's *topic-quiz*. If a student received equally low grades on more than one quiz in the same section (e.g., 76% on Q1 and Q3), they were able to choose between the topics. An illustration of the submission timeline for the course assessments and example grades can be seen in Fig. 1. Students completed Q1–Q6 in the third, fifth, eighth, 10th, 12th, and 14th weeks of the 14-week semester, respectively. The students began working on their midterm projects after they received all their quiz grades for the first half of the semester but did not turn in the midterm project until after they completed quiz 4 (in week 11). DYOP final projects were due during the final exam period approximately 1.5 weeks after quiz 6.

Quiz Assessment. Students took six quizzes throughout the semester. Each quiz followed the same general format: two-to-three multiple-choice questions and one-to-two free-response questions. The multiple-choice questions were sourced from the Fluid Mechanics Concept Inventory [62,63]. These questions were chosen because they had been designed specifically to assess a student's conceptual understanding of key topics in fluid mechanics. The free-response questions focused more heavily on a student's ability to interpret a problem statement and appropriately apply a calculation-based solution strategy. An example quiz can be found in Ref. [64].

DYOP Assessment. The DYOP assignment is an Authentic Learning Assignment in that it is designed to begin transitioning

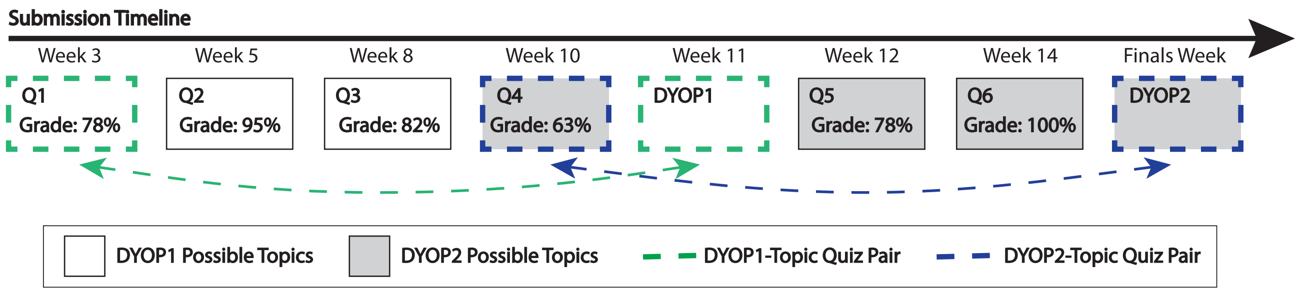

Fig. 1 Diagram of assessment workflow throughout the semester. An illustration of how a student would select their topic for their DYOP assignment is given. Numbers indicating grades are provided as an illustration only and do not come from a student who participated in this study.

students toward engaging with subject matter material in a manner authentic to practicing engineers. This assignment asks students to research a real-life topic that can be modeled leveraging one or more of the strategies used in class. Students are tasked with researching the application, modeling the problem (which requires them to research fluid properties, equipment specifications, and system requirements), and then solving the problem. This type of assignment is also commonly referred to as *Student Generated Questions* [65]. An example of the assignment prompt used for the DYOP can be found in Ref. [64]. While the DYOP assessment prompt itself was design-based, key components of this assessment include re-engagement with course material (an opportunity not typically provided in traditional lecture-based classrooms), extended time and open resources, student-driven problem framing, and a clear rubric for evaluation. Each key feature is reflective of the iterative engineering design process. We emphasize that not only the assessment design but also its implementation and integration into the course are key components of this intervention.

Reflection Assignment. While reflection assignments in engineering might span behavior-based and cognitive outcomes, this reflection prompt was developed with the goal of having students reflect specifically on their solution strategies (a focus on cognition). The prompt was developed over the course of three months in a collaborative faculty development seminar on reflections [66]. The prompt asks students to describe their problem-solving process and identify the areas that challenged them. In this way, students are encouraged to engage with the material for a second time in a low-pressure environment. One key component of this reflection assignment is that students were asked to reflect on their submissions prior to receiving the solutions to the quiz. Therefore, students had to use their best judgment and conceptual understanding of the material to analyze and reflect on their work. The reflection prompt asked questions such as *How did you decide on a solution strategy for this problem?* and *What assumptions did you make while solving the problem? How?* The full prompt can be found in Ref. [59]. The reflection prompt was graded for completion-only. To promote honest reflections, students were told that the instructor would not read the reflections until after the semester had ended. Students were required to complete reflections within 24 h following the completion of the quiz, before corresponding quiz solutions or grades were released. DYOP reflections were submitted with the DYOP. Reflection prompts were kept consistent across the quiz and DYOP reflections.

Research Design

The current Institutional Review Board (IRB)-approved study explores the extent to which students engage with subject matter material according to Bloom's Taxonomy in a typical assessment compared to that of an authentic learning assignment. The following questions guided our research:

- How does the novel authentic learning assignment, DYOP, affect engagement at upper levels of learning (as defined by Bloom's Taxonomy) when compared to traditional assessments?
- Given that significant differences are found between student cognitive engagement on the quiz and DYOP assessments, how might these differ as a function of race/ethnicity, gender, prior experiential learning experiences, and initial quiz performance?

Validity Considerations. This was a single-section, instructor-embedded study. Using another section as a control would have confounded instructor and section effects. With only one section available, we prioritized internal coherence (same instructor, same students) over a nonequivalent comparison. We used written reflections to measure expressed cognitive engagement, recognizing that reflection can itself prompt metacognition. In prior work [59], we estimated reflection-induced content and found that for remember–analyze, most statements captured cognition that occurred during the quiz rather than being generated by the act of reflecting; only evaluate and N/A were predominantly reflection-induced (though statistically significant differences were not found). Because the reflection prompt, length, and completion-only grading were identical across conditions (quiz and DYOP), we expect any reflection effect to operate similarly across both, reducing differential bias. Even so, if reflections inflate quiz-level codes such as evaluate and N/A, this would bias our current study *against* detecting differences and making any observed quiz–DYOP increases at these levels conservative. To limit demand characteristics, students were not told the study purpose and were told that the instructor would not read their reflections (completion checked only), reducing incentives for students to "perform" higher-order cognition in their reflection responses.

Participants. Students were recruited to participate at the beginning of the semester by a member of the research team who was not involved in the instruction of the course. As all elements of the study were integrated into the course design, student participation in the study required no extra labor. As such, only one of the students in the 55-person class was ineligible to participate in the study, leaving 54 eligible students who elected to participate. As seen in Table 1, of the eligible students, 37 identified as a man, 16 identified as a woman, and one identified as nonbinary. The majority of the class, 31 students, were white, and 28 students identified as students of color and more than one race. Additionally, nearly half of the students participated in an internship or Co-Op in previous semesters. Specific details of the demographic data that were collected can be seen in Table 1.

While there were 54 students who participated in this study, each student submitted two topic-quizzes and corresponding DYOP assignment reflections that can be included for analysis (e.g., 108 total submissions). Only submissions for which both the topic-quiz and corresponding DYOP had complete reflections were included in the analysis, resulting in a total of $N = 97$ topic-quiz/DYOP reflection pairs available for analysis.

Instrument. To address the questions laid out in this study, content analysis of student reflections was conducted. Using Bloom's taxonomy as a conceptual framework, a codebook was developed for a priori deductive coding. The course-instructor and a student researcher (who had excellent undergraduate-level proficiency in fluid mechanics) first created a codebook based on learning objectives from the course. The codebook had six codes corresponding to the different levels of Bloom's Taxonomy: *remember*, *understand*, *apply*, *analyze*, *evaluate*, and *create*. Then they worked together to expand the codebook to operationalize the different ways that the various Bloom's levels might manifest during execution of the course assessments. Following the development of the initial codebook, five student reflections were selected at random and coded. The raters then met to discuss any disagreements across applied codes and refine the codebook further. This process was iterated on until consensus was reached that the codebook was complete. It was during this process that a seventh code, *not applicable (N/A)*, was developed. Sentences coded as *N/A* included statements that were often focused on the student's emotional state or material related more to the experience of taking the quiz rather than the quiz content itself. The complete codebook can be found in the Appendix (Table 7) and Ref. [64], with extensive examples of student reflections coded at each level found in Ref. [59].

To ultimately test our data for statistically significant differences in engagement at each Bloom's level, the reflection data need to be "quantized"—transforming the qualitative reflection data into numerical data [67,68]. Therefore, it was decided that for each reflection, a calculation of the proportional coverage for each Bloom's level would be calculated (details regarding this calculation are in the following section). To ensure consistency across reflections, heuristics for how codes would be applied were developed. Thus, the reflections were coded at each level of Bloom's Taxonomy in complete sentences such that each sentence could be coded as one or more of the Bloom's Taxonomy levels. Applying codes to complete sentences provided a means of precise and consistent application of code coverage, but necessitated that codes were not mutually exclusive from one another as students often mentioned multiple levels in one sentence. Some sentences might have clauses that encompass themes from multiple levels of Bloom's simultaneously, while others may have two clauses that represent different Bloom's taxonomy levels. For example, a sentence can be coded as both the *remember* and *understand* levels of Bloom's Taxonomy. An example of a sentence coded as both *understand (indicated in **bold text**)* and *remember* (indicated in underlined text) is shown below:

> *For the second part of the question, **I saw that it was a find the force acting on the hinged door type of problem**, so I decided that following the method in our notes (outlined in Module 2) was the best way to go about doing it.*

In the example above, the student understood the prompt of the question and the important characteristics, so the first independent clause is highlighted blue to indicate this student is at the *understand* level of Bloom's Taxonomy. The second independent clause is the student describing their decision to refer to their notes from class which includes a set of equations and definitions, so this section is highlighted yellow indicating this student is in the *remember* level of Bloom's Taxonomy. Therefore, the whole sentence is coded as *understand* and *remember*. The only code that is considered mutually exclusive with the various Bloom's levels is the *not applicable (N/A)* code. Thus, it was ensured that the representative "amount" of code applied was consistent across reflections and interraters.

To further ensure quality in the quantization of the qualitative data, interrater reliability and Cohen's κ were used to evaluate the strength of interrater agreement. Codes were applied using qualitative analysis software NVIVO. After the first researcher coded the reflections, a second researcher coded a random sample of 25% of the reflections for each assessment. The κ values for each assessment are shown in Table 2 categorized by Bloom's taxonomy level in each column. The values in the table range from moderate (0.41–0.6) to near-perfect (0.81–1) agreement, with half showing substantial agreement (0.61–0.8). For this study, a minimum acceptable κ value of 0.41 was set for each quiz-code pair average across reflections, and a minimum acceptable value of 0.61 (substantial agreement) was set for the complete reflection across all codes. If interrater agreement fell below this threshold, raters met and discussed differences until consensus was reached. Due to pragmatic constraints around interrater time and the sheer amount of coding required in this analysis (648 total reflections—162 interrated), moderate minimum agreement thresholds for quiz-code pairs were selected. Furthermore, because Cohen's κ is sensitive to prevalence and bias (e.g., sparse categories) and we expect variation in coverage to result from the variation in quiz topics (i.e., Reynolds transport theorem applications require greater levels of analysis than Bernoulli problems), we report Cohen's κ by level and percent agreement for each assessment to better contextualize reliability.

Quantization of Data Into Percent-Coverage Values. Once reflections had been coded for each level of Bloom's, the relative level of student engagement at each level is quantified by calculating the percent coverage of each code (%Coverage = # of words coded per Bloom'slevel/total # of words in reflection). For example, out of a student's entire reflection, 30% of the reflection might be coded as *remember* while only 5% is coded as *analyze*. The result is a transformed set of qualitative data that is now represented as a quantitative set of continuous data

Table 1 Demographic information

Individual-level variables	N	Percent (%)
Gender		
Woman	16	29.6
Man	37	68.5
Non-binary	1	1.9
Age in years		
17–19	11	20.4
20–22	43	79.6
Ethnicity/race[a]		
Black or African American	4	7.4
Asian, Native Hawaiian, or other Pacific Islander	15	27.8
White	31	57.4
Hispanic or Latino	5	9.3
More than one race	3	5.6
Other	1	1.9
Major		
Mechanical Engineering	49	90.7
Nuclear and Radiological Engineering	4	7.4
Computer Engineering	1	1.9
Year of undergraduate study		
2	18	33.3
3	27	50.0
4	9	16.7
Internship/co-op experience prior to taking class		
Yes	23	42.6
No	31	57.4

[a]Students were given the option to select all that apply. Totals do not equal 100%.

Table 2 Cohen's κ agreement for Bloom's Taxonomy codes

	Cohen's κ							
	Remember	Understand	Apply	Analyze	Evaluate	Create	N/A	Total agreement
Quiz 1	0.69	0.49	0.61	0.61	0.71	1	0.81	0.71
Quiz 2	0.62	0.55	0.7	0.67	0.92	1	0.75	0.74
Quiz 3	0.79	0.73	0.8	0.46	0.97	1	0.92	0.81
Quiz 4	0.54	0.5	0.68	0.5	1	1	0.88	0.72
Quiz 5	0.46	0.51	0.66	0.44	1	1	0.85	0.70
Quiz 6	0.81	0.75	0.83	1	1	1	0.93	0.90
DYOP 1	0.81	0.63	0.86	0.92	0.94	0.74	0.67	0.79
DYOP 2	0.91	0.97	0.97	0.97	0.99	0.97	0.94	0.96

ranging from 0% to 100%. This quantification allowed us to better analyze larger trends in the engagement level across different assessments and student groupings. This method is a form of summative content analysis [69]. Accordingly, a greater percent coverage is taken to indicate higher priority by the writer: a larger percent coverage of a given Bloom's level means the writer placed greater emphasis on that cognitive level in their reflection (and in the assessed work) than on other levels. This process, through the utilization of a categorization framework, facilitated the use of statistical tests without generalizing the population [70].

The percent coverage is calculated using the word count (as opposed to character count). A comparison between word count and character count was conducted, and it was found that the difference in percent coverage calculated between the two methods was less than 1%. This is approximately equivalent to 10 words which is well within the uncertainty levels at each code based upon the interrater agreement. Therefore, the word count metric was used for this calculation as it was less cumbersome to extract from the NVIVO coding software. The reader should note that the percent coverage for each reflection can sum to greater than 100%. This is due to the fact that codes at each Bloom's level are not mutually exclusive. This analysis allows us to calculate the percent coverage at each level of Bloom's Taxonomy for each student.

Statistical Analysis. To answer the aforementioned research questions, this study conducted statistical analyses to determine the percent-coverage difference between the students' topic-quiz reflections and DYOP reflections. Analyses were planned as within-person contrasts of DYOP versus quiz for Bloom coverage. Our primary interest was investigating outcomes in higher-order categories (analyze, evaluate, and create) based on the intervention's theory of change; tests on other categories presented in this article are considered exploratory. It should also be noted that because reflections were constrained to one-page, Bloom-level percent coverage is compositional. Thus, percent-coverage differences reported here reflect the relative emphasis students placed on different cognitive processes within a fixed narrative space. Additional analyses were conducted to evaluate the effect of topic-quiz grade, demographics, and work experience; these are also constructed as exploratory queries.

To address our primary research question, *How does the DYOP assignment affect engagement at various Bloom's Taxonomy levels of learning when compared to traditional assessments?*, a standard t-test was conducted on each data stream (*remember*, *understand*, *apply*, *analyze*, *evaluate*, *create*, and *N/A*). The t-test allows us to compare the percent-coverage means of the two groups (1. topic-quiz reflection coverage versus 2. DYOP reflection coverage) and determine if they are statistically significantly different. Notably, the coverage data spread for each taxonomy level are different; some code coverages have a normal, nonskewed spread with no outliers, while others have a very skewed, non-normal spread with outliers. Assumptions for paired t-tests were evaluated on DYOP–quiz difference scores using Shapiro–Wilk (SW) and Kolmogorov–Smirnov tests. Differences for remember and understand did not deviate from normality (SW $p > .10$), while apply, analyze, evaluate, create, and N/A showed statistically detectable departures (SW $p < .05$). With $N \approx 97$, few outliers were identified (0–3 per category). Fortunately, the sample size for this data is sufficiently large ($N = 97$) and therefore the central limit theorem can be imposed such that the t-test is used for analysis. To verify the efficacy of this assumption, the t-test, Wilcoxon signed-rank test, and sign tests were run (as applicable based on assumption criteria). These resulted in consistent conclusions throughout, and therefore only the t-test results are presented here. For all tests, a 95% confidence interval ($\alpha = 0.05$) was used.

Our secondary research question, posed as an exploratory question, focuses on how the DYOP assignment might differently affect different groups as a function of demographic variables like race/ethnicity and gender, topic-quiz performance, and experiential learning experiences (internships and co-ops). To conduct this analysis, a second variable was generated that represents the percent-coverage difference for each code:

$$\%\text{Coverage Difference}_{\text{DYOP-TQ}} = \text{DYOP \% Coverage} - \text{Topic Quiz \% Coverage}$$

For example, if the *analyze* level comprised 30% coverage on a DYOP reflection and 10% coverage on the corresponding topic-quiz reflection, then the percent-coverage difference would be 20%. Thus, tests for differences between groups like analysis of variance (ANOVA) and the Mann–Whitney U test can be employed. While this variable can be useful for determining if certain groups benefit more or less from the DYOP (a greater benefit represented by a larger %Coverage difference$_{\text{DYOP-TQ}}$), it is a function of two independent variables. Therefore, any significant differences found across groups could originate from one variable versus the other. To provide insight into the source of these differences, for any test on %Coverage difference$_{\text{DYOP-TQ}}$, a parallel test for each of the independent % coverage variables was run.

The first between-group test assesses the growth in a student's engagement level based on their initial topic-quiz grade. We seek to answer questions such as whether a student who received an F on their topic-quiz showed more improvement in Bloom's level engagement during DYOP compared to a student who received a B. Thus, grade groups are defined. The dataset is divided into groups based on topic-quiz grades, which averaged 67% (not unreasonably low given that the topic-quiz is, by definition, the student's worst quiz grade). Due to the skewed distribution toward low grades, a K-means clustering analysis [71] was employed to determine unique clusters. Fortunately, this resulted in relatively evenly distributed grade groups ($n = 29$, 35, and 33, respectively) with grade divisions at 65.5% and 79.1%. Thus, an A/B, C/D, and D/F group was generated. DYOP grades were not considered for grade-grouping purposes due to minimal variation (mostly A grades). Comparisons between groups were made by running ANOVAs on all three variables. For any between-group

Table 3 DYOP–topic-quiz percent-coverage differences by Bloom's level

	Difference							
	Minimum	Mean	Maximum	σ	Skewness	Kurtosis	t	p-Value
Remember	−31.92	4.0	45.21	2.9	0.41	0.68	2.9	0.005
Understand	−50.25	−16.7	41.20	−9.1	0.73	1.10	−9.1	<0.001
Apply[a]	−56.75	−3.5	33.24	−1.9	−0.11	0.16	−1.9	0.05
Analyze[b]	−33.50	5.7	30.25	6.4	−0.14	3.81	6.4	<0.001
Evaluate[a]	−7.10	13	37.56	12.3	0.38	−0.79	12.3	<0.001
Create[a]	0.00	26.6	75.23	17.3	0.65	0.32	17.3	<0.001
NA[a]	−61.46	−20.6	29.50	−5.0	−0.16	−0.70	−5.0	<0.001

Notes: Values show descriptive statistics for the difference scores, standard deviation (σ), skewness, and kurtosis, along with inferential tests comparing DYOP versus topic-quiz. Results from the paired t-tests (two-tailed, $\alpha = .05$) are shown; $N = 97$ paired reflections.
When Shapiro–Wilk indicated non-normal differences, confirmatory nonparametric tests were run and are marked as follows: [a]sign test, [b]Wilcoxon signed-rank.

comparison that indicated significant differences, Tukey's honestly significant difference (HSD) was conducted.

Next, we were interested in whether the DYOP impacted cognitive growth as a function of work experience status, gender-minority status, and race/ethnicity-minority status. First student submissions were grouped by student-relevant work experience (e.g., having participated in an engineering internship or co-op previously). This resulted in ($n = 45$) samples from students who had experience and ($n = 52$) samples from those without. Gender differences were examined by dividing student submissions into two

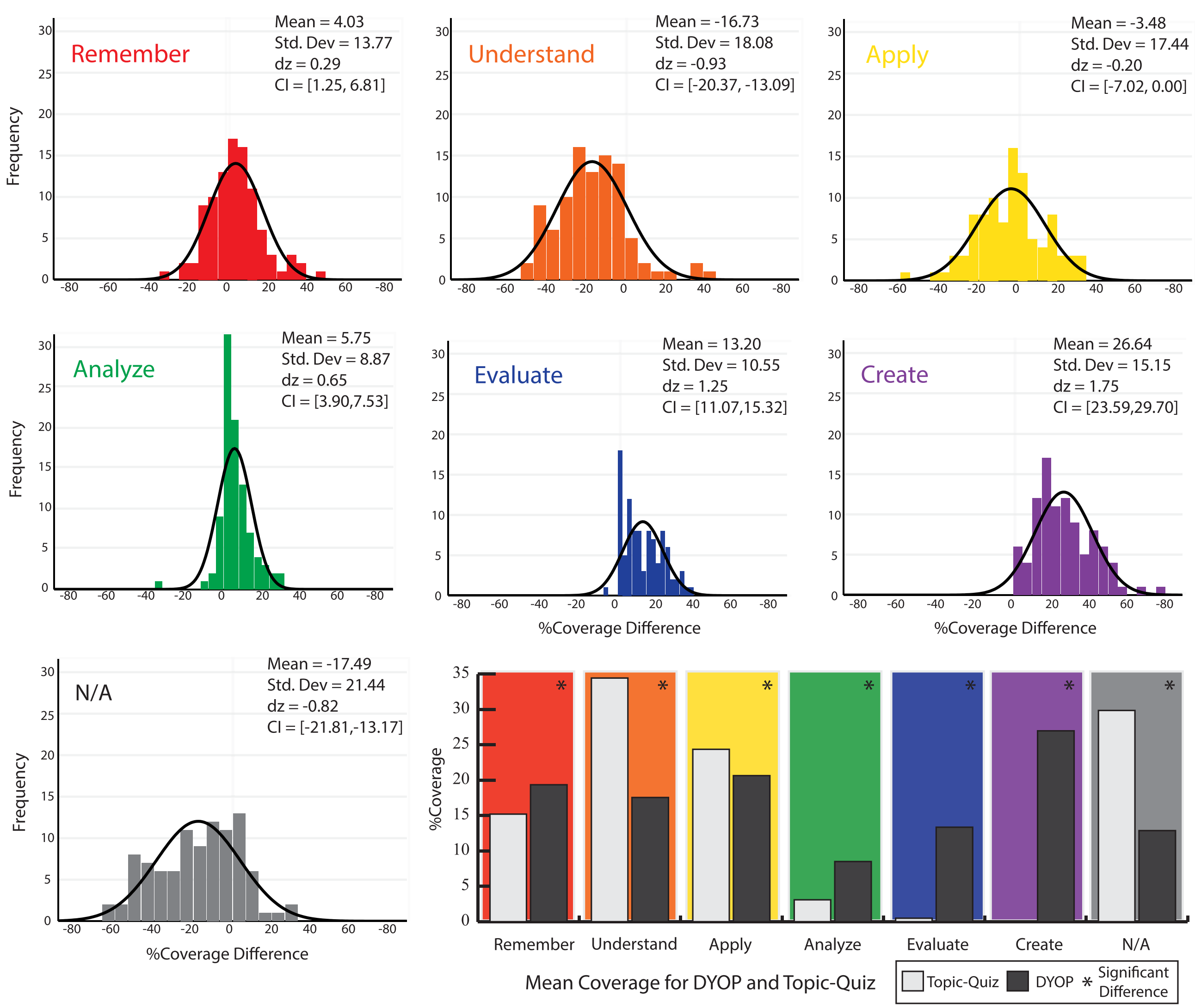

Fig. 2 Histograms: % coverage difference between the DYOP and topic-quiz reflections for each level of Bloom's. For each, the mean, standard deviation (std. dev.), effect size (dz), and 95% confidence interval have been included. Bar chart: mean % coverage for each level of Bloom's for both the DYOP and topic-quiz reflections.

Table 4 One-way ANOVA results comparing Bloom-level percent coverage across grade groups

	% coverage difference$_{DYOP-TQ}$		% coverage DYOP		% coverage topic-quiz	
	F	p-Value	F	p-Value	F	p-Value
Remember	1.54	0.22	1.12	0.33	0.36	0.70
Understand	0.10	0.90	0.42	0.67	0.91	0.41
Apply	0.44	0.64	0.71	0.49	0.41	0.67
Analyze	9.23	<0.001*	9.73	<0.001*	4.28	0.02*
Evaluate	2.35	0.10	1.71	0.19	1.30	0.28
Create	0.09	0.92	0.09	0.92	–	–
NA	0.59	0.59	4.11	0.02*	1.17	0.31

For each Bloom's level, we tested between-group differences on three outcomes: (1) the DYOP–topic-quiz coverage difference (percentage points), (2) DYOP coverage, and (3) topic-quiz coverage. Values are F statistics with two-tailed p-values ($\alpha = .05$). Grade groups were defined by final course grade bands with divisions at 65.5% and 79.1%. $N = 97$. Results are interpreted cautiously given multiple tests.

Table 5 Tukey's HSD post hoc comparisons among grade bands at the analyze level following the significant one-way ANOVAs in Table 4

	% coverage difference$_{(DYOP-TQ)}$			% coverage DYOP			% coverage topic-quiz		
Mean diff.	A/B group	C/D group	D/F group	A/B group	C/D group	D/F group	A/B group	C/D group	D/F group
A/B group	–	8.7*	5.3*	–	7.42*	8.04*	–	−1.34	2.57
C/D group	–	–	−3.4	–	–	0.62	–	–	3.91*
D/F group	–	–	–	–	–	–	–	–	–

Notes: Outcomes are percentage-point measures for (1) DYOP–topic-quiz coverage difference, (2) DYOP coverage, and (3) topic-quiz coverage. Entries are mean differences computed as row−column; positive values indicate higher coverage in the row group. Significant contrasts (familywise $\alpha = .05$ within each ANOVA) are marked with*.
Group divisions: (A/B) 79.2–100%, (C/D) 65.6–79.1%, and (D/F) 0–65.5%.

groups: women/nonbinary ($n = 30$) and men ($n = 67$). This distribution aligns with the gender-minority status in the engineering program at this institution. Race/ethnicity-based groups comprised of student submissions from nonwhite students ($n = 48$) and white ($n = 49$) students. The nonwhite group includes all students who selected "Black or African American," "Asian or Pacific Islander," "Hispanic/Latinx," or "other" due to low numbers in each individual minority group. The Mann–Whitney U test, suitable for nonparametric data with independent groups and no assumption of homogeneity of variance, was employed for these analyses. The independent t-test was not used due to the lack of homogeneity of variance and insufficient group sizes to impose the central limit theorem.

Limitations

This study evaluated the DYOP in a single, core mechanical engineering course grounded in Newtonian mechanics; generalizability to other domains should be tested. The implementation occurred during remote instruction in the COVID-19 period, which may have influenced results, although similar patterns have since been observed in in-person offerings (outside the scope of this IRB). Reflections served both as the measurement of expressed cognition and as a metacognitive component of the quizzes and DYOP package, so some measurement–intervention interaction is possible; prior work from our group found no significant reflection-induced differences at higher Bloom levels [59], but reactivity cannot be ruled out. The one-page reflection cap also makes percent coverage a compositional measure, so our estimates capture relative emphasis in students' discourse rather than total cognition. It should also be noted that Bloom's Taxonomy is hierarchical, thus to engage at higher levels, one must pass through lower levels, so increases in upper Bloom's levels (which may correspond to less discussion of lower levels during reflection) do not necessarily indicate less engagement at those levels during assessment. Moreover, repeated engagement with the same topic (quiz then DYOP) likely contributes to learning; without a control group, we cannot isolate the effect of any single component of the DYOP package. Finally, we conducted paired comparisons for seven Bloom categories: analyze, evaluate, and create were specified a priori as primary outcomes, with other levels treated as exploratory. We did not apply multiplicity corrections and therefore interpret marginal p-values cautiously; future work testing this or similar interventions should preregister outcomes, include controlled or counterbalanced designs, and control the false discovery rate.

Results and Discussion

DYOP Impact on Bloom's Taxonomy Level Engagement. Table 3 shows the descriptive statistics, skewness, kurtosis, and

Table 6 Tukey's HSD results for differences found between grade groups at the N/A level

	% coverage DYOP		
Mean diff.	A/B group	C/D group	D/F group
A/B group	–	−6.10	−7.54*
C/D group	–	–	−1.44
D/F group	–	–	–

Notes: Tukey's HSD post hoc comparisons among grade bands at the N/A level following the significant one-way ANOVA in Table 4 (DYOP coverage only). Entries are mean differences in percentage points (row–column); positive values indicate higher N/A coverage for the row group. Significant p-values (familywise $\alpha = .05$ within this ANOVA) are marked with an asterisk (*).
Group divisions: (A/B) 79.2–100%, (C/D) 65.6–79.1%, and (D/F) 0–65.5%.

results of the *t*-test for the percent-coverage comparison between the topic-quiz and DYOP reflections for each Bloom's Taxonomy level. For all levels of Bloom's Taxonomy, the mean value of the DYOP reflection coverage was significantly different from the DYOP topic-quiz reflection coverage. *Understand*, *analyze*, *evaluate*, *create*, and N/A have $p < 0.001$; *remember* has $p = 0.005$; and *apply* has $p = 0.05$. To better interpret these results, Fig. 2 shows histograms of the % coverage difference variable for each Bloom's Taxonomy level. Positive values indicate higher DYOP percent coverage compared to the topic-quiz percent coverage for the respective Bloom's Taxonomy level, while negative values indicate the opposite.

It can be seen in Fig. 2 that *remember*, *analyze*, *evaluate*, and *create* Bloom's levels show a higher percent coverage in the DYOP reflection compared to that of the topic-quiz reflection. It can also be seen that *understand*, *apply*, and *N/A* Bloom's levels show lower percent coverage in the DYOP reflection compared to that of the topic-quiz reflection. These results might be explained by the fact that students were requested to write a 1-page reflection for both the topic-quiz and DYOP. Given that the total length of the reflections is consistent, it is reasonable that for some levels to increase, others must decrease. Since problems designed by students on the DYOP were of similar complexity as those they completed on quizzes, a decrease in percent coverage at the *understand* and *apply* levels likely indicates less emphasis placed on engagement at these levels during reflection rather than less engagement at them. Indeed, it would be difficult for students to engage at the *analyze* level without first having engaged fully in the *understand* and *apply* levels. The decrease in *N/A* (−20.6%) may reflect lower performance-related commentary (e.g., anxiety) during DYOP compared to quizzes.

This highlights the unintended yet positive result of the DYOP intervention: students are provided an opportunity to demonstrate mastery of a course topic in a format that de-emphasizes the testing anxiety felt by most students, allowing them to focus more on course content and their own learning. In general, the higher levels of Bloom's Taxonomy showed an increase in percent coverage in the DYOP assignment, with a mean percent-coverage increase of 13% for *evaluate* and 26.6% for *create.* These results suggest the DYOP assignment was associated with deeper engagement, particularly at higher Bloom's levels, in this course offering.

Influence of Topic-Quiz Grade on Cognitive Growth. Table 4 shows the results of the ANOVA that tested differences between grade groups. The grade groups broadly represent students who made A/B grades (group 1), C/D grades (group 2), and D/F grades (group 3). For each Bloom's level, three separate ANOVAs were run: 1. on the % coverage difference$_{(\mathrm{DYOP\text{-}TQ})}$, 2. on the % coverage DYOP, and 3. on the % coverage topic-quiz. Table 4 includes the *F*-statistic, with higher *F*-values indicating that the variance between grade groups is larger than the variance within grade groups, and a *p*-value <0.05 indicating there are significant differences found between grade groups. As our main focus is on cognitive growth seen between the topic-quiz and DYOP, we will first focus on the ANOVA results from the % coverage difference$_{(\mathrm{DYOP\text{-}TQ})}$ variable. The only significant difference found between grade groups was at the *analyze* level. Interestingly, when looking at the ANOVA results for the % coverage DYOP and % coverage topic-quiz variables, a significant difference was found between grade groups at this level for both variables. This indicates that the growth seen at the *analyze* level is influenced by both independent variables. To better understand the nature of this difference in growth between grade groups, the results of Tukey's HSD for the *analyze* ANOVAs can be seen in Table 5. First, for the percent-coverage difference between the DYOP reflection and topic-quiz reflection, group 1 is significantly different than groups 2 and 3 with $p < 0.001$ and $p = 0.039$, respectively.

A similar result was found for the ANOVA run on the percent coverage for the DYOP reflections only. The A/B group is significantly different than the C/D and D/F groups with $p < 0.001$ for both. Last, for the ANOVA run on the percentcoverage for the topic-quiz reflections, only a significant difference was found between the C/D and D/F groups with $p = 0.01$. When looking at the DYOP reflections alone, these results indicate that the A/B group engaged at the *analyze* level more than the C/D and D/F groups. This could indicate a possible higher level of complexity of the problems designed by the A/B students for their DYOP assignment. Further analysis of the DYOP submissions would be required to determine if this is the case. Alternatively, when looking at the topic-quiz reflections alone, these results indicate that the C/D group engaged at the *analyze* level more than both the A/B and D/F group while taking their quiz. This result indicates that there might exist an optimum amount of uncertainty regarding the accuracy of an answer or solution strategy that will prompt students to engage more deeply at the *analyze* level. It follows that this optimum value would exist in the middle grade group.

Finally, a significant difference was found between groups for the ANOVA run on the % coverage DYOP variable at the *N/A* level. A significant difference was found between the A/B and D/F groups. In this case, the D/F group engaged at the *N/A* level more than the A/B group (as shown in Table 6). Since the *N/A* comments were representative of generalized comments on the assessment and statements representing student confidence, anxiety, and uncertainty, it makes sense that the D/F group might face the most uncertainty and anxiety while completing the DYOP compared to the other groups. The unbounded nature of the DYOP requires a reasonable grasp of fundamentals for students to even begin to identify a problem or system to model that relates to the course content. It is likely that the D/F group struggled to identify and scope a problem more than the students in the C/D and A/B groups, resulting in higher levels of anxiety felt during the completion of this project.

Influence of Demographic Factors on Cognitive Growth. We conducted exploratory Mann–Whitney U comparisons of Bloom-level coverage between demographic groups (co-op versus no co-op; male versus nonmale; white versus nonwhite) for both topic-quiz coverage and the DYOP–quiz coverage difference. Across tests, we found no statistically significant differences, with one exception: a difference at the analyze level for the DYOP–quiz coverage change between male and nonmale students ($p = 0.04$). Given the number of tests, the compositional nature of the outcomes, and our aggregated groupings, this isolated finding should be interpreted cautiously. Accordingly, we feel these exploratory results indicate a potential for the DYOP intervention to promote comparable engagement across broad groups thus promoting equity; however, aggregation of minority categories can mask heterogeneity. Determining whether DYOP promotes equity requires a study specifically designed and powered for that purpose (e.g., disaggregated categories, preregistered primary outcomes, and error-rate control). Here, we simply note the potential of the DYOP assessment to promote equity in engineering and leave quantification of this impact to future work.

Conclusion

This study reveals that an authentic learning assignment like the DYOP can increase student engagement with subject matter material at the higher levels of cognitive learning. By incorporating an assessment that requires the application of design-based skills such as problem framing and definition to solve an open-ended problem (without a predetermined solution), students are guided to more deeply engage with subject matter concepts. Exploratory analyses did not show consistent differences across broad demographic or experience groups, which may indicate the potential for equitable engagement; targeted studies would be needed to confirm this.

The analysis of percent-coverage comparison between topic-quiz and DYOP reflections across Bloom's Taxonomy levels reveals significant differences. While *understand*, *apply*, and *N/A* levels show lower percent coverage in the DYOP reflection compared to the topic-quiz reflection, *remember*, *analyze*, *evaluate*, and *create* levels demonstrate higher coverage in the DYOP reflection. These findings suggest a trade-off between different cognitive levels during reflection writing, possibly influenced by the nature of the DYOP assignment and decreased performance anxiety. Notably, higher Bloom's Taxonomy levels exhibit substantial percent-coverage increases in DYOP reflection, emphasizing the effectiveness of DYOP assignments in fostering deeper engagement, particularly at advanced cognitive levels.

The ANOVA analysis revealed insights into the engagement levels of students across different grade groups, focusing on their engagement levels in DYOP reflections, topic-quiz reflections, and the difference between the two. Overall, there were no significant differences found between grade groups, suggesting similar levels of engagement on both assessments across all levels of Bloom's Taxonomy irrespective of quiz accuracy. Students achieving near-perfect grades and those struggling with mastery benefited similarly from the DYOP. However, notable exceptions were observed at the *analyze* level, where group 1 (A/B grades) displayed significantly higher engagement compared to groups 2 and 3 on their DYOP reflections. This disparity indicates potential differences in problem complexity designed by students. Conversely, in topic-quiz reflections, group 2 (C/D grades) showed greater *analyze*-level engagement than groups 1 and 3 (A/B and D/F grades), suggesting an optimal level of uncertainty during assessment might prompt deeper analysis. Additionally, a significant difference was noted at the *N/A* level for the DYOP reflections, where group 3 (D/F grades) exhibited higher engagement, possibly influenced by student feelings of anxiety and lack of confidence.

Ultimately, this study presents a feasible approach for integrating design concepts into a core engineering course. This integration is a first step toward breaking the dichotomy between convergent and divergent thinking in engineering training. Not only does integrating design concepts into core-curriculum courses improves the depth of cognitive engagement with technical material, but it has the potential to facilitate more equitable pathways toward concept mastery (as demonstrated by consistently high grades on DYOP assessments and higher cognitive engagement across all demographic groups). Strategies such as these could be considered for entry- and mid-level courses, with the potential to support both design skills and technical content mastery.

Acknowledgment

We would like to acknowledge the efforts of undergraduate researchers, Ticiano Gao, Angelina Rodriguez, and Comforter Afobunor who supported this study by serving as raters and conducting literature review for this study.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

Appendix

Table 7 Coding guide for all levels of Bloom's Taxonomy

Code	Indicators
Remember	• Student cites/states facts/definitions, memorized equations not in the context of the way they are solving the problem • Student references equations • Student notices a mistake (either during the test or during reflection) but does not correct the mistake
Understand	• Student demonstrates understanding of how an equation/fact is to be used • Assumption: student makes assumption that comes directly from the problem statement • Student demonstrates ability to understand important characteristics of the problem/student restates or summarizes problem in their own words • Student can demonstrate incorrect understanding, but still be applying understanding as it makes sense to them • Student understands what certain equations mean and the context of the current problem they are solving
Apply	• Free body diagram (FBD): student completes FBD as part of a process they are repeating • Assumption: student assumes from practice/applies a correct assumption but does not demonstrate reasoning behind it • Student applies skill they have practiced before • Student uses an equation to define a system • Student solves equations even when stating they are unsure • Student solves or implies that they solved equation • Student states or restates answer: emphasis on the action of solving the problem • Student catches mistake and recalculates equation/describes correct answer
Analyze	• Student checks/defends answer using different assumption or solution method, but does not provide an assessment of the impact of their decision or different solution • Free body diagram: student provides reasoning/logic behind why FBD is used • Assumption: student makes assumption that is not directly given in the problem statement AND provides a defense based on physical understanding of why assumption applies • Student identifies multiple ways to arrive at an answer, chooses one, and defends it • Student responds to new information—must indicate not having seen before, no similar examples, etc.—by analyzing an approach to solve problem • Student explains link between an equation and its application/impact • Catching mistakes: when a student has a reason to go back and check their answer or do the problem again another way
Evaluate	• Student provides reasoning for certainty/uncertainty of applicability or accuracy of their solution in the context of an engineering problem • Student evaluates the efficacy of their solution (ex: checking answers)
Create	• Student provides insight into problem design/mentions combining different principles of fluid mechanics or engineering to design a problem • Student explains reasoning behind new method to solve problem that was not taught in class • Student discusses process of modeling a real-world engineering problem within the scope of a typical course example problem • Student recognizes an initial solution/model was flawed and makes adjustments to better represent the engineering problem (revision of original concept/problem) • Student develops the model of the problem • Student ideates or brainstorms • Student independently develops or demonstrates a skill that is novel to them while doing this assignment
N/A	• General comments on the problem • Statements of confidence with no technical reasoning/support • Generic statements of certainty/uncertainty • Comments on testing strategy

References

[1] Selter, J. J. A., 2012, "Infusing Professional Skills Development Into Co-op Student," 2012 ASEE Annual Conference & Exposition, San Antonio, TX, June 10–13.

[2] Kousar Basha, P., Suneetha, Y., Chithra, V. B., and Kiran Kumar Reddy, G., 2024, "Identification of Soft Skills Gaps in Engineering Education: A Theoretical Study," Bull. Technol. Hist. J., **24**(7), pp. 100–107.

[3] Lattuca, L. R., Terenzini, P. T., Volkwein, J. F., and Peterson, G. D., 2006, "The Changing Face of Engineering Education," Bridge-Washington-Natl. Acad. Eng., **36**.

[4] Abbatiello, A., Agarwal, D., Bersin, J., Lahiri, G., Schwartz, J., and Volini, E., 2018, "The Rise of the Social Enterprise," Deloitte Global Human Capital Trends.

[5] Mourshed, M., Farrell, D., and Barton, D., 2012, "Education to Employment: Designing a System That Works," mckinseyonsociety.com/education-to-employment, Accessed October 15, 2025.

[6] Brunhaver, S. R., Korte, R. F., Barley, S. R., and Sheppard, S. D., 2018, "Bridging the Gaps Between Engineering Education and Practice," *US Engineering in a Global Economy*, National Bureau of Economic Research, University of Chicago Press, Chicago, IL, pp. 129–164.

[7] Casper, A. M. A., Atadero, R. A., Hedayati-Mehdiabadi, A., and Baker, D. W., 2021, "Linking Engineering Students' Professional Identity Development to Diversity and Working Inclusively in Technical Courses," J. Civ. Eng. Educ., **147**(4), p. 04021012.

[8] Pfeiffer, F. M., Burgoyne, S., Hunt, H. K., and Strobel, J., 2018, "Balancing Academic Rigor and Creative Thinking: A Transformational Approach to Teaching Senior Design," ASME J. Biomech. Eng., **140**(8), p. 081011.

[9] Therriault, D. J., Douglas, E. P., Buten, E., Bates, E. A., Berry, M. B., and Waisome, J. A. M., 2025, "Are There Problems With Academic Problems? Examining Problem Types in Two Statics Engineering Textbooks," Eur. J. Eng. Educ., **50**(4), pp. 764–779.

[10] Clancy, S. M., 2024, *Examining Underemphasized Engineering Practices: Engineering Students' Alignment With Curricular Emphases and Engagement With Divergent Thinking*, University of Michigan, Ann Arbor, MI.

[11] Lin, W. L., and Lien, Y. W., 2013, "Exploration of the Relationships Between Retrieval-Induced Forgetting Effects With Open-Ended Versus Closed-Ended Creative Problem Solving," Thinking Skills and Creativity, **2013**(10), pp. 40–49.

[12] Editorial, G., 2011, "What Constitutes a Good Design Education Research Paper That Would Be Suitable for JMD?," ASME J. Mech. Des., **133**(3), p. 030301.

[13] Gavin, K., 2011, "Case Study of a Project-Based Learning Course in Civil Engineering Design," Eur. J. Eng. Educ., **36**(6), pp. 547–558.

[14] Neal, M. T., 2013, *Development, Evaluation, and Impact of a Hands-On Introductory Course in Mechanical Engineering*, The Ohio State University, Columbus, OH.

[15] Shekar, A., 2014, "Project-Based Learning in Engineering Design Education: Sharing Best Practices," 2014 ASEE Annual Conference & Exposition, Indianapolis, IN, June 15–18.

[16] Shuhailo, Y. V., and Derkach, T. M., 2021, "Project-Based Learning for Undergraduate Engineering Students Minoring in Textile Technology and Design," J. Phys. Conf. Ser., **1840**(1), p. 012042.

[17] Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., and Eagan, M. K., 2018, "Anatomy of STEM Teaching in North American Universities," Science, **359**(6383), pp. 1468–1470.

[18] Lombardi, M. M., and Oblinger, D., 2007, "Approaches That Work: How Authentic Learning Is Transforming Higher Education," EDUCAUSE Learning Initiative (ELI) Paper 5.

[19] Rule, A. C., 2006, "The Components of Authentic Learning," J. Authent. Learn., **3**(1).

[20] Bloom, B., 1956, *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*, David McKay Company, New York.

[21] Krathwohl, D. R., 2002, "A Revision of Bloom's Taxonomy: An Overview," Theory Pract., **41**(4), pp. 212–218.

[22] Cross, N., 2001, *Engineering Design Methods: Strategies for Product Design*, 4th ed., Wiley, Hoboken, NJ.

[23] Dorst, K., and Cross, N., 2001, "Creativity in the Design Process: Co-evolution of Problem-Solution," Des. Stud., **22**(5), pp. 425–437.

[24] Khorshidi, M., Shah, J. J., and Woodward, J., 2014, "Applied Tests of Design Skills—Part III: Abstract Reasoning," ASME J. Mech. Des., **136**(10), p. 101101.

[25] Fink, L. D., 2013, *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*, John Wiley & Sons, San Francisco, CA.

[26] Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., and Zhang, D., 2008, "Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage 1 Meta-Analysis," Rev. Educ. Res., **78**(4), pp. 1102–1134.

[27] Halpern, D. F., 2003, "Teaching for Critical Thinking: Helping College Students Develop the Skills and Dispositions of a Critical Thinker," New Dir. Teach. Learn., **1999**(80), pp. 69–74.

[28] Kazerounian, K., and Foley, S., 2007, "Barriers to Creativity in Engineering Education: A Study of Instructors and Students Perceptions," ASME J. Mech. Des., **129**(7), pp. 761–768.

[29] Deitrick, A. R., and Berdanier, C. G. P., 2024, "Exploring Engineering Graduate Students' Perceptions of Creativity in Academic and Research Environments," 2024 ASEE Annual Conference Proceedings, Portland, OR, June 23–26.

[30] Peng, A., Menold, J., and Miller, S. R., 2023, "Nature Versus Nurture: The Influence of Classroom Creative Climate on Risk-Taking Preferences of Engineering Students," International Design Engineering Technical Conferences and Computers an Information in Engineering Conference, Boston, MA, Aug. 20–23.

[31] Beaulieu, D. F., 2022, "Creativity in Science, Engineering, and the Arts: A Study of Undergraduate Students' Perceptions," J. Creat., **32**(3), p. 100035.

[32] Beghetto, R. A., and Kaufman, J. C., 2016, *Nurturing Creativity in the Classroom*, Cambridge University Press, Cambridge, UK.

[33] Field, B. W., 2007, "Visualization, Intuition, and Mathematics Metrics as Predictors of Undergraduate Engineering Design Performance," ASME J. Mech. Des., **129**(7), pp. 735–743.

[34] Girgis, M., 2010, "A New Engineering Taxonomy for Assessing Conceptual And Problem Solving Competencies," 2010 Annual Conference & Exposition, American Society of Engineering Education, Louisville, KY, June 20–23.

[35] Deo, S., and Hölttä-Otto, K., 2024, "Critical Thinking Assessment in Engineering Education: A Scopus-Based Literature Review," ASME J. Mech. Des., **146**(7), p. 072301.

[36] Leonard, A., Guanes, G., and Dringenberg, E., 2023, "Undergraduate Students' Beliefs About Diverse Approaches to Making Engineering Design Decisions: Exploring Change During a Capstone Course," Int. J. Technol. Des. Educ., **33**(5), pp. 1959–1989.

[37] Barry, B. E., Ohland, M. W., Mumford, K. J., and Long, R. A., 2016, "Influence of Job Market Conditions on Engineering Cooperative Education Participation," J. Prof. Issues Eng. Educ. Pract., **142**(3), p. 04015017.

[38] Ammerman, R., and Sen, P., 2006, "An Undergraduate Power Engineering Curriculum: A Unique and Practical Approach to Bridging the Gap Between Academia and Industry," 2006 Annual Conference & Exposition, Chicago, IL, June 18–21.

[39] Baburajan, V., Silva, J. D. A. E., and Pereira, F. C., 2021, "Open-Ended Versus Closed-Ended Responses: A Comparison Study Using Topic Modeling and Factor Analysis," IEEE Trans. Intell. Transp. Syst., **22**(4), pp. 2123–2132.

[40] Septiani, S., Retnawati, H., and Arliani, E., 2022, "Designing Closed-Ended Questions Into Open-Ended Questions to Support Student's Creative Thinking Skills and Mathematical Communication Skills," JTAM, **6**(3), pp. 616–628.

[41] Robertson, B. F., Walther, J., and Radcliffe, D. F., 2007, "Creativity and the Use of CAD Tools: Lessons for Engineering Design Education From Industry," ASME J. Mech. Des., **129**(7), pp. 753–760.

[42] Marra, R. M., Palmer, B., and Litzinger, T. A., 2000, "The Effects of a First-Year Engineering Design Course on Student Intellectual Development as Measured by the Perry Scheme," J. Eng. Educ., **89**(1), pp. 39–45.

[43] Vaidya, A., 2021, "Contributions to the Teaching and Learning of Fluid Mechanics," Fluids, **6**(8), p. 269.

[44] Bondehagen, D. L., 2011, "Inspiring Students to Learn Fluid Mechanics Through Engagement With Real World Problems," 2011 ASEE Annual Conference & Exposition, Vancouver, BC, Canada, June 26–29.

[45] Albers, L., and Bottomley, L., 2011, "The Impact of Activity Based Learning, a New Instructional Method, in an Existing Mechanical Engineering Curriculum for Fluid Mechanics," 2011 ASEE Annual Conference & Exposition, Vancouver, BC, Canada, June 26–29.

[46] Kapilan, N., Vidhya, P., and Gau, X. Z., 2021, "Virtual Laboratory: A Boon to the Mechanical Engineering Education During Covid-19 Pandemic," High. Educ. Future, **8**(1), pp. 31–46.

[47] Mandavgane, S., 2020, "Fun With Fluid: An Innovative Assignment in Fluid Mechanics," Educ. Chem. Eng., **2020**(30), pp. 40–48.

[48] Webster, D. R., Kadel, R. S., and Newstetter, W. C., 2020, "What Do We Gain by a Blended Classroom? A Comparative Study of Student Performance and Perceptions in a Fluid Mechanics Course," Int. J. Eng. Educ., **36**(1(A)), pp. 2–17.

[49] Kaiphanliam, K. M., Nazempour, A., Golter, P. B., Van Wie, B. J., and Adesope, O. O., 2021, "Efficiently Assessing Hands-On Learning in Fluid Mechanics at Varied Bloom's Taxonomy Levels," Int. J. Eng. Educ., **37**(3), pp. 624–639.

[50] Setiyana, R., and Muna, H., 2019, "Evaluating Bloom's Taxonomy in Test Items," Proceedings of the English Education International Conference, Banda Aceh, Indonesia, Sept. 18–19, pp. 79–83.

[51] Chandio, M. T., Pandhiani, S. M., and Iqbal, R., 2016, "Bloom's Taxonomy: Improving Assessment and Teaching-Learning Process," J. Educ. Educ. Dev., **3**(2), pp. 203–221.

[52] Kadiyala, S., Gavini, S., Kumar, D. S., Kiranmayi, V., and Rao, P. V. L. N. S., 2017, "Applying Blooms Taxonomy in Framing MCQs: An Innovative Method for Formative Assessment in Medical Students," J. Dr. NTR Univ. Health Sci., **6**(2), pp. 86–91.

[53] Dewey, J., 1933, *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*, DC Heath & Co, Boston, MA.

[54] Rodgers, C., 2002, "Defining Reflection: Another Look at John Dewey and Reflective Thinking," Teach. Coll. Rec., **104**(4), pp. 842–866.

[55] Baruah, B., Ward, T., and Jackson, N., 2017, "Is Reflective Writing an Effective Peer Assessment Tool for Students in Higher Education?" 2017 16th International Conference on Information Technology Based Higher Education and Training (ITHET), Ohrid, Macedonia, July 10–12.

[56] Williams, J. M., 2002, "The Engineering Portfolio: Communication, Reflection, and Student Learning Outcomes Assessment," Int. J. Eng. Educ., **18**(2), pp. 199–207.

[57] Kember, D., McKay, J., Sinclair, K., and Wong, F. K. Y., 2008, "A Four-Category Scheme for Coding and Assessing the Level of Reflection in Written Work," Assess. Eval. High. Educ., **33**(4), pp. 369–379.

[58] Evans, E., Menold, J., and McComb, C., 2019, “Critical Thinking in the Design Classroom: An Analysis of Student Design Reflections,” International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Anaheim, CA, Aug. 18–21.
[59] Yraguen, B., Lummus, A., Koolman, H., Moore, R., and Fu, K., 2022, “Using Post-assessment Reflection to Enhance Student Learning Outcomes in a Fluid Mechanics Course,” 2022 ASEE Annual Conference & Exposition, Minneapolis, MN, June 26–29.
[60] Sparks-Langer, G. M., Simmons, J. M., Pasch, M., Colton, A., and Starko, A., 1990, “Reflective Pedagogical Thinking: How Can We Promote It and Measure It?,” J. Teach. Educ., **41**(5), pp. 23–32.
[61] Kember, D., 2008, *Reflective Teaching and Learning in the Health Professions: Action Research in Professional Education*, John Wiley & Sons, Hoboken, NJ.
[62] Martin, J., Mitchell, J., and Newell, T., 2003, “Development of a Concept Inventory for Fluid Mechanics,” 33rd Annual Frontiers in Education, Boulder, CO, Nov. 5–8.
[63] Martin, J. K., Mitchell, J., and Newell, T., 2004, “Work in Progress: Analysis of Reliability of the Fluid Mechanics Concept Inventory,” Proceedings—Frontiers in Education Conference, Savannah, GA, Oct. 20–23, IEEE.
[64] Yraguen, B. F., Koolman, H. E., Lummus, A. M., Fu, K. K., and Moore, R. A., 2024, “Promoting Equity and Cognitive Growth: The Influence of an Authentic Learning Assignment on Engineering Problem-Solving Skills,” 2024 Annual Conference & Exposition, American Society of Engineering Education, Portland, OR, June 23–26.
[65] Angelo, T. A., and Cross, K. P., 2012, *Classroom Assessment Techniques*, Jossey Bass Wiley, San Francisco, CA.
[66] Mejia, K. Z., Turns, J. A., Flores, Y., Ritz, H., Liao, J., Chen, J., and Yraguen, B. F., 2021, “Benefits of Codesigning With Educators as Faculty Development,” Paper Presented at the 2021 ASEE Virtual Annual Conference, Virtual Conference, July 26–29.
[67] Sandelowski, M., and Barroso, J., 2003, “Creating Metasummaries of Qualitative Findings,” Nurs. Res., **52**(4), pp. 226–233.
[68] Sandelowski, M., 2003, “Tables or Tableaux. The Challenges of Writing and Reading Mixed Methods Studies,” *Handbook of Mixed Methods in Social and Behavioral Research*, A. Tashakkori and C. Teddlie, eds., Sage Publications, Inc., Thousand Oaks, CA, pp. 321–350.
[69] Hsieh, H. F., and Shannon, S. E., 2005, “Three Approaches to Qualitative Content Analysis,” Qual. Health Res., **15**(9), pp. 1277–1288.
[70] Borrego, M., Douglas, E. P., and Amelink, C. T., 2009, “Quantitative, Qualitative, and Mixed Research Methods in Engineering Education,” J. Eng. Educ., **98**(1), pp. 53–66.
[71] Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., and Heming, J., 2023, “K-Means Clustering Algorithms: A Comprehensive Review, Variants Analysis, and Advances in the Era of Big Data,” Inf. Sci., **2023**(622), pp. 178–210.