2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #50852

Work in Progress: Preliminary Characterization of Analysis, Evaluation, and Testing Activities in Mechanical Engineering Capstone Design Reports

Nadine El-Ghaffir, University of Michigan

Nadine El-Ghaffir is a graduate student at the University of Michigan pursuing a Master's degree in Mechanical Engineering and a Ph.D. in Design Science. Her research interests include product development and design processes, design methods and methodologies, and engineering design education. She received her B.S. in Chemical Engineering from the University of Michigan.

Dr. Shanna R. Daly, University of Michigan

Shanna Daly is an Associate Professor in Mechanical Engineering at the University of Michigan. She has a B.E. in Chemical Engineering from the University of Dayton and a Ph.D. in Engineering Education from Purdue University.

Dr. Kathleen H. Sienko, University of Michigan

Kathleen H. Sienko is an Arthur F. Thurnau Professor of Mechanical Engineering at the University of Michigan (UM). She earned her Ph.D. in 2007 in Medical Engineering and Bioastronautics from the Harvard-MIT Division of Health Science and Technology, and

Work in Progress: Preliminary Characterization of Analysis, Evaluation, and Testing Activities in Mechanical Engineering Capstone Design Reports

Abstract

Engineering design requires making and justifying evidence-based decisions throughout a design process, from early problem framing through validation. Preparing students to engage in these decisions is a central goal of engineering design education; however, how novice designers enact key decision-support activities – analysis, evaluation, and testing – across design phases remains insufficiently understood. This paper presents a qualitative document analysis of capstone design reports from senior-level undergraduate engineering teams to characterize how these activities are enacted across design phases. Analysis is defined as the use of analytical or computational models to examine concept performance; evaluation as the assessment of concepts against metrics, objectives, and specifications; and testing as the examination of how concepts perform in practice [1]. Findings indicate that these activities are infrequently documented during front-end design phases and are instead concentrated in the middle and later stages of the design process. Across reports, analysis and evaluation activities were most often used to support concept comparison and down-selection, frequently through decision matrices such as Pugh charts, while testing was primarily framed around confirming or validating selected concepts rather than informing early design direction. Importantly, there appears to be a missed opportunity to more directly connect analysis and testing activities to the criteria and evidence used within evaluation and decision matrices during concept selection. Taken together, these patterns suggest that novice designers tend to position analysis, evaluation, and testing as downstream selection and verification and validation mechanisms rather than as integrated, iterative activities for early-stage exploration and framing. The findings highlight the need for instructional approaches that more explicitly support the early and continuous use of analysis, evaluation, and testing to strengthen evidence-based design decision-making across all phases of a design process.

Introduction

Engineering design requires designers to make and justify evidence-based decisions throughout design processes. These decisions are critical to project success, as they shape the quality, cost, and desirability of the resulting solution [2]. However, without structured approaches to generating and using evidence, designers may rely on intuition or incomplete information, leading to biased or inconsistent decision-making (e.g., preference effects or self-contradictory choices) [2].

Analysis, evaluation, and testing are core practices that support more rigorous and defensible decision-making across a design process. Analysis enables designers to use analytical or computational models to anticipate performance, identify uncertainties, and guide subsequent inquiry [1], [3]. Evaluation supports systematic comparison among alternatives by applying metrics, objectives, and criteria, helping to surface assumptions and clarify trade-offs [4]. Testing provides empirical evidence about how concepts perform in practice, informing iteration and reducing the risk of downstream failure [5], [6]. Given their central role in supporting evidence-based decision-making under uncertainty, it is essential to equip designers with the tools and training needed to effectively engage in analysis, evaluation, and testing throughout a design process.

Existing literature points to differences in how experienced designers and novice designers approach evaluation and testing. For example, a study examining differences in how novice and experienced designers approach design tasks within the aerospace industry suggested that experienced designers tend to evaluate their decisions prior to implementation, whereas novice designers tended to rely on "trial and error" patterns, resulting in several iterations and a longer cyclic process [7]. In that study, novice designers tended to generate, implement, and then evaluate; in contrast, experienced designers tended to generate, conduct a preliminary evaluation, implement, and then conduct a final evaluation. Complementing these findings, prior work indicates that novice designers frequently make feature changes through unintentional testing. The same study found that physical models provide immediate feedback that helps correct inaccurate mental models and increases the proportion of functional ideas [8]. Collectively, this evidence suggests that novice designers' approaches are comparatively less structured and intentional than those of experienced designers.

Despite extensive attention to the role of analysis, evaluation, and testing in design, there is limited understanding of how these activities are enacted by novice designers, particularly students, across different phases of a design process. To address this gap, we conducted a qualitative document analysis of final reports from a senior-level mechanical engineering capstone design course, consistent with prior engineering education studies that analyzed team/capstone project documents (e.g., [9], [10]). For this exploratory study, we identified and characterized student teams' analysis, evaluation, and testing activities across three phases of design processes: the front-end, middle, and the back-end phases. The objective was to provide an initial characterization of how student teams perceive and conduct those activities during their capstone design projects, laying the foundation for future research as well as the development of tools or instructional supports aimed at strengthening students' use of analysis, evaluation, and testing throughout their design work.

Background

Analysis, Evaluation, and Testing Definitions

There are differences in how analysis, evaluation, and testing are defined and discussed across engineering disciplines and application domains. This variability is reflected in engineering design education. For instance, Cross described analysis as an activity that is often done numerically or analytically [4]; Pahl et al. described analysis using examples such as weak-spot, fault-tree, and reliability analyses, where the goal is to assess risks, safety, tolerances, and failures [11]; and Dieter and Schmidt described analysis as a way to understand the expected behaviors of a concept's components before it exists in a physical form by using appropriate disciplines of science and the necessary computational tools [12]. In addition, definitions for testing vary significantly. For example, Tahera et al. reported that testing has been defined in electrical engineering as "an activity in which a system or component is executed under specified conditions," where "the results are observed or recorded, and an evaluation is made of some aspect of the system or component" [13]. On the other hand, Banks described testing as an activity not limited to functional testing; instead, activities such as administering reviews, inspections, and walkthroughs are "similar to devising a test under which model accuracy is judged" [14]. These variations are emphasized by Patel at al., who investigated the current state of testing based on design textbooks and noted that "reviewing the selected textbooks through a lens of testing in engineering design revealed that none of the books provide an explicit definition of the term testing" [15].

Analysis, Evaluation, and Testing Across Design Phases

Beyond textbook definitions, prior work has examined how analysis, evaluation, and testing are implemented across different different phases of a design process. Dym et al. emphasized the role of these three activities throughout the embodiment, and detailed design phases [1]. Regarding analysis, Dym et al. described its role during embodiment design, where the proposed concept from the conceptual phase is further modeled and analyzed [1]. These analytical models provide designers with critical information regarding performance of a concept before any additional resources are committed to a detailed design. In terms of evaluation, Dym et al. suggested that concepts are generated and evaluated against functional specifications to choose a concept after comparing different alternatives during the conceptual design phase [1], which aligns with prior work highlighting the role of evaluation during concept generation and selection [16]. For example, Burchill described a systematic review of generated concepts during ideation and suggested that concept selection involves a systematic review, comparison, and enhancement of concepts in an iterative process [16]. The author further suggested that individual and group thinking, expert feedback, and experiments in the laboratory may be used iteratively to combine and improve upon initial solution concepts, developing a small set of promising concepts which are then evaluated in detail against customer requirements and organizational constraints in order to select the dominant concept(s) [16]. With regard to testing, Tahera et al. suggested that testing is described as an activity that occurs throughout a design process, ranging from component-level testing to proof-of-concept testing and prototype development [13]. For example, during the detailed design phase, the final design of a concept is articulated in greater detail, where the concept from the preliminary design is refined and optimized. Consistent with this perspective, Dym emphasized testing preliminary designs or evaluating them against metrics to determine whether concepts satisfy design objectives [17];

therefore, the author proposed a "generate and test" strategy in which an automated generation strategy is used to create a large set of solutions that are then tested against relevant metrics and constraints, thus enabling the generation or search for a solution to be redirected based on test results. The authors also described a back-tracking strategy to rectify the problem by undoing the decisions that contributed to a failure [17]. Overall, the role of testing in the later stages of design was discussed, noting that detailed or final designs may require revisiting decisions made earlier in a process.

Activities Performed by Novice Designers

Existing literature offers a view into how novice designers perform analysis, evaluation, and testing, highlighting the strategies and tools they use to support design decisions. A study of student teams' decision-making processes found that all but one team used Pugh charts to support down-selection, possibly reflecting course requirements or prior familiarity with this method [18]. The same study also reported that students engaged in a range of social decision-making strategies and structured methods, including preference voting, spectrum voting, individual and group Pugh charts, multiple-round elimination, single-leader decision-making, and group consensus. In terms of testing, prior research consistently emphasized the benefits of continuous, hands-on testing across multiple phases of a design process. For example, research on senior design projects highlighted the value of a "design–build–test" approach, in which iterative prototyping and testing supported students' understanding of a well-structured design process and enabled integration across disciplines [19]. Supporting this, a study involving first-year college students found that systems tested at multiple stages of development demonstrated higher reliability and improved performance [20]. A parallel can be seen in software design through the "test-driven development" (TDD) framework, where tests are used to guide system design [21]. TDD emphasizes a shift from traditional "test-last" approaches to a "test-first" model, positioning testing as a continuous and integral component of the development process rather than a final step [22], [23].

Despite the recognized importance of analysis, evaluation, and testing in supporting evidence-based design decisions, there remains limited understanding of how novice designers engage in these activities across the full design lifecycle. In particular, the distribution of these activities across early problem framing, concept development, and later verification and validation in student design work remains undercharacterized.

Methods

This qualitative document analysis examined how student teams reported engaging in analysis, evaluation, and testing across the front-end, middle, and back-end phases of their design work. The study was guided by the following research questions:

1. *What analysis, evaluation, and testing activities do student engineering design teams perform during capstone design processes?*
2. *How do student teams' analysis, evaluation, and testing activities differ across phases of a design process?*

Data Collection, Participants, and Context

Nine publicly available reports from a senior-level mechanical engineering capstone design course at a large Midwestern university were examined. The number of reports was deemed appropriate for a document analysis, as qualitative research guides emphasize purposeful selection over the selection of a large number of documents [24]. Methodological guidance and review literature further suggest that sample size in such studies depends on the research questions and reaching data saturation [25], [26].

All nine reports were produced by senior-level undergraduate mechanical engineering student teams and were systematically selected to represent a range of project domains and course enrollment semesters, spanning Winter 2022 through Fall 2025. The projects covered a variety of engineering design project areas, such as nanotechnology and materials, automotive, marine, consumer products, energy, robotics, mechanical and structural engineering, and industrial and manufacturing. Since analysis, evaluation, and testing practices varied by project type, the sample was intentionally drawn across domains to capture a range of approaches and reduce redundancy associated with a single project type. All selected reports involved product-oriented design projects, where teams developed physical solutions that typically culminated in full-scale builds or high-fidelity prototypes demonstrating key aspects of form and/or function.This shared context provided a consistent basis for comparing teams' analysis, evaluation, and testing activities across projects. These projects are summarized in Table 1, including semester of study, project domain, and report number.

Table 1. Details of the reports reviewed (semester, domain, and report number)

Semester	Domain	Report Number
Winter 2022	Consumer Products	Report 9 [27]
Fall 2023	Industrial & Manufacturing	Report 3 [28]
Winter 2023	Nanotechnology & Materials Science	Report 6 [29]
Fall 2024	Mechanical & Structural engineering	Report 8 [30]
Fall 2024	Automotive	Report 4 [31]
Fall 2024	Marine Technology	Report 5 [32]
Winter 2024	Energy	Report 1 [33]

Winter 2025	Marine Technology	Report 2 [34]
Winter 2025	Robotics	Report 7 [35]

The course typically includes five to six sections of four to six student teams, each led by a different instructor. Projects are generally sponsored by external or internal partners, with whom student teams collaborate in defining authentic engineering problems and developing corresponding design concepts. Throughout the design course, student teams are required to conduct analysis, evaluation, and testing as part of a design process. However, teams retain autonomy in determining the specific types of analyses, evaluations, and tests to perform, as well as the level of detail and number of times these activities are conducted. These decisions are made by student teams based on the needs and constraints of their individual projects, resulting in variation in approaches across teams and project domains. Instruction is provided at the section level rather than through centralized lectures, and may vary across instructors; however, all sections maintain consistent pedagogical goals and require completion of the same core learning activities. For example, all students are required to complete standardized learning blocks and achieve the same milestones. Relevant learning blocks include the "Engineering Analysis and Testing Block," and "Verification and Validation Block." The Engineering Analysis and Testing Block covers topics such as the role of analysis and testing in a design process, the questions each method answers, modes of testing, usability testing, design for "X," optimization, and systems thinking. Through this learning block, students are encouraged to identify "worries" in a concept and to support their concept development decisions with evidence. Within this block, risk and safety are also covered, including topics such as risk assessment and Failure Mode and Effects Analysis (FMEA). The Verification and Validation Block highlights the definitions of verification and validation (V&V), the questions they answer, their role in early and later stages of a design process, and the ways V&V can be applied to the course projects. The learning block further exposes students to real-world V&V examples and tracing matrices, emphasizing the importance of organization and accountability. The learning blocks primarily consist of readings, and are supplemented with additional instructional resources such as videos to provide students with further support.

Over the course of the semester, student teams complete four written milestones: Design Review 1 (DR1), Design Review 2 (DR2), Design Review 3 (DR3), and the Final Report. DR1 requires teams to define the problem statement and establish requirements and specifications. DR2 requires teams to generate ideas, down-select to their top concept(s), and conduct initial development and iteration. DR3 requires teams to (1) perform and document appropriate analyses and tests to address their concept "worries"; (2) use results to iterate and refine their concept; (3) provide a build description that systematically captures decisions related to sourcing, pricing, acquisition, and fabrication/manufacturing to demonstrate the feasibility and completeness of the proposed solution; (4) present the final concept while linking design decisions to analysis and testing results; and (5) propose verification and validation plans detailing and justifying methods for evaluating performance. The Final Report requires student teams to revise and integrate prior content from DR1, DR2, and DR3, and to report details about the final design in addition to verification and validation results.

Data Analysis

Following the selection of reports for inclusion in this study, a deductive-inductive analysis was conducted, resulting in a codebook that captured students' analysis, evaluation, and testing activities during their design work. Our data analysis approach was drawn from Walther's Procedural Validation, which entails systematic and documented processes of analysis and interpretation [36]. This meant that each report was reviewed multiple times, with each review focused on a specific analytical goal, such as the categorization of activities. At each stage of the analysis, we documented report data and interpretations, creating an auditable and traceable record.

The focus of the analysis was to identify how student teams approached analysis, evaluation, and testing during their design work. Guiding this analysis were the definitions of analysis, evaluation, and testing from Dym, Little, and Orwin's engineering design textbook [1]. Analysis was defined as an analytical or computational model used to "simulate its performance under a stated set of conditions"; evaluation as an activity in which designers assess a chosen concept against metrics that reflect the client's objectives and specifications that stipulate how the concept must function; and testing as an activity used to obtain data about how well a concept actually works [1].

To conduct the deductive analysis, these definitions were used in combination with an initial list of possible activities constructed by our research team. This list was developed based on prior research focused on novice designer prototyping practices (e.g., [36]) and our team's expertise and experience teaching the course. This deductive list included analytical models, computational models, selection methods, experiments, and user testing. Using this list, we reviewed design activities described in the reports and coded instances as analysis, evaluation, or testing when they were (1) explicitly reported by student teams, (2) consistent with the study's definitions, and (3) included sufficient detail about what was assessed and how the assessment was conducted. We examined the presence of these activities across each report and recorded a binary indicator (yes/no) for whether a given type of analysis, evaluation, or testing was evident. We did not count the frequency of individual instances; instead, our goal was to indicate whether each student team, as a whole, demonstrated engagement with each type of activity. Table 1 presents the coding schemes for analysis, evaluation, and testing in the design reports. In addition, we performed inductive analysis to capture emerging themes related to how these activities were described in the reports.

Table 1. Description of deductive activities

Category	**Subcategory**	**Subcategory Definition**
Analysis	Computational models	Running software-based calculations, models, or simulations
	Analytical models	Running manual calculations using basic, established, or advanced equations
Evaluation	Selection methods	Evaluating concepts against a set of design constraints
Testing	Experiments	Conducting controlled tests to determine how changes in variables affect performance or behavior
	User testing	Engaging representative users with a concept to assess how well it supports their goals, needs, and usability

As a second layer of analysis, we differentiated design activities based on their timing within a design process: the front-end phase, middle phase, and back-end phase. Table 2 describes the focus of each phase and provides examples illustrating the types of activities considered in each.

Table 2. Phase definitions with illustrative examples

Phase	**Definition**	**Illustrative Activities Used to Guide Analysis**
Front-end	This phase focused on design activities associated with problem identification and definition, requirements and specifications development, background research, and concept generation	- Rapid experimental tests using low-fidelity prototypes to assess conceptual feasibility - Preliminary calculations or modeling to assess feasibility of key requirements
Middle	This phase focused on activities associated with concept development, concept selection, and	- Physical prototypes to assess performance during the concept down-selection process

	feasibility assessment of concept solutions	- Computational models or simulation to compare performance of concept alternatives
Back-end	This phase focused on design activities associated with verification and validation	- Final prototype testing under realistic operating conditions - Validation testing with users to assess usability and functionality

Findings

Analysis of the student team capstone reports revealed frequent documentation of computational models, analytical models, and experiments during activities corresponding to the DR3 deliverables in the middle phase. In addition, reports showed that decision matrices, such as Pugh charts, were most often used during concept down-selection by student teams. Evidence of computational models and experiments was also apparent in the back-end phase. Overall, the reports indicated limited analysis, evaluation, and testing in the front-end phase of the design projects, followed by frequent engagement with these activities during the middle phase, and a continued focus on analysis and testing during the back-end phase. Table 3 summarizes the analysis, evaluation, and testing activities evidenced across the nine capstone reports by design phase.

Table 3. Analysis, evaluation, and test activities described in design reports. Values indicate the number of reports (N = 9) containing evidence of each activity

Category	**Student Analysis, Evaluation, and Testing Activities**	**Reports with evidence of Activity**	**Reports with Evidence in Front-End Phases**	**Reports with Evidence in Middle Phases**	**Reports with Evidence in Back-End Phases**
Analysis	*Computational models*	9	0	7	6
	Analytical models	7	0	7	3
Evaluation	*Selection methods*	9	0	9	0

Testing	*Experiments*	9	1	6	7
	User testing	3	0	0	3

<u>*Computational Models*</u>
The capstone reports documented many computational models within the design work, with seven of the reports including these models during the middle phase to: (1) determine and confirm dimensions, (2) check and compare hand calculations, (3) determine feasibility, (4) ensure components met expected performance, and (5) ensure components met requirements. For example, Report 7 indicated the use of a Python script and NumPy library to verify that the concept dimensions allowed the reach of the required force during the middle phase of design. The report described it as follows:

> *This analysis uses the Jacobian matrix of the finger, and returns the grasp force on the object of all fingers [19]. The calculation considers 3D space of the fingers (accounting for all directions), and it is conducted in python, using numpy library. (Report 7, Middle Phase)*

In the back-end phase, three reports documented the use of analytical models to: (1) verify components would not fail, (2) verify specifications, and (3) ensure functionality of concepts. For example, Report 3 included analytical models to verify that the concept's components would not fail under a specified set of conditions during the back-end phase of design. The report described it as follows:

> *Approach: Use CAD software to determine mass of components, calculate weight, compare system weight to average bottlenose dolphin weight, including analysis relative to average female and male bottlenose dolphin weights. Evaluate for build design and final design." (Report 5, Back-End Phase)*

None of the reports documented the use of computational models during the front-end phase.

<u>*Analytical Models*</u>
The capstone reports documented various analytical models within the design work, with seven of the reports indicating the use of analytical models during the middle phase. In these instances, reports indicated the use of analytical models to: (1) determine unknown parameters, (2) identify key variables affecting the performance of the design or its components, (3) ensure components met specifications, (4) justify selection of design components, (5) predict performance against requirements and specifications, (6) establish operational requirements and steps for future users, and (7) assess risk and safety. For example, Report 8 described the use of analytical models to

verify that the materials would not fail under expected stress from the vehicle during the middle phase of design. The report described it as follows:

> *We verified that the PLA will be able to support the stress from the ride vehicle, but also need to make sure that our MDF support will not break under the forces being imposed on it. To do this, we use the conservation of momentum and centripetal acceleration equations from above, as well as the moment equation for beam bending." (Report 8, Middle Phase)*

In the back-end phase, three reports documented the use of analytical models to: (1) verify components would not fail, (2) verify specifications, and (3) ensure functionality of concepts. For example, Report 3 included analytical models to verify that the concept's components would not fail under a specified set of conditions during the back-end phase of design. The report described it as follows:

> *We performed fatigue life calculations to verify that the pinion teeth would not fail within the specified lifetime of 2190 complete cycles. Assuming that each width adjustment was the entire span of the range(18"), this means that pinion teeth would come into contact with the rack teeth at most 18 times per cycle. Thus, the pinion teeth would need to endure 39,420 total cycles of a compressive stress of 85.5 MPa (the calculated Hertz stress)." (Report 3, Back-End Phase)*

Risk, safety, and reliability were also addressed through analytical models. For example, Report 5 identified a dangerous scenario, chose a maximum operating condition, performed a simulation to estimate the maximum system pressure, and compared results to a defined maximum safety threshold. The report described it as follows:

> *We further sought to confirm the safety of the Venturi tube dimensions in performing the simulation assuming the constant flow of water at the maximum dolphin swim speed of 15m/s into the tube, which would make for the maximum pull obtainable by our system. As shown in the resultant simulation in Figure 10, the system would not impose a pressure with a magnitude greater than our maximum safety pressure of 5.87 psi. This overall confirmed the dimensions of our Venturi tube design in concept, which we then constructed in our build design to perform physical testing.* (Report 5, Middle Phase)

While the majority of analytical models were used during the middle and back-end phases, none of the reports described their use during the front-end phase.

<u>*Selection Methods*</u>

The use of concept selection methods, such as decision matrices, was documented only during the middle phase, which focused on concept development and selection. All nine reports documented the consistent use of concept selection methods, including Pugh charts, to guide the down-selection process. Notably, reports indicated that student teams often employed these methods as a stand-alone approach to choose their selected concept, rather than pairing them with iterative analysis or testing to validate assigned scores. Reports also showed that concepts with the highest numerical scores were typically prioritized and selected by student teams. In addition, the relative meaning of assigned scores was seldom articulated, making it unclear whether small differences in ratings corresponded to meaningful differences between concepts. For example, Report 5 noted: “Totalling all of the scores the two concepts with the highest point values became our alpha and beta designs.”

Across reports, there was little to no evidence that analysis or testing activities were used to generate or substantiate the scores assigned within these decision matrices. For example, six of nine reports documented reliance on subjective judgment and perceived feasibility during down-selection. In several cases, reports included outlines of advantages and disadvantages of concepts; nevertheless, the highest-scoring concept was ultimately selected. This is exemplified by Report 2:

> *Concept 4 also did not score the highest in the waterproof category, which was judged based on the manufacturing method for the design. Stereolithography (SLA) 3D printed resin (for Concept 4) is not as waterproof as molded silicone (for Concept 3); however, 3D printing is a simpler, more accessible, and better for prototyping manufacturing method. Both the resin and silicone are waterproof, but we determined that silicone's greater advantage in waterproofing did not outweigh resin's manufacturing advantage.* (Report 2, Middle-Phase)

<u>*Experiments*</u>

All reports documented experiments during the design process. During the front-end phase, only one report (Report 6) documented experiments used for initial familiarization with equipment and processes. The report described it as follows:

> *The goal of our initial testing was to get hands on with the grinder itself and the capillaries, and to understand the scale and precision of the task our team was assigned.”* (Report 6, Front-End Phase)

Experiments were frequently documented in the middle phase. In this phase, reports described their use to: (1) determine optimal parameters, (2) determine unknowns, (3) confirm functionality of components, and (4) assess design effectiveness. For example, Report 9 indicated the use of experiments to determine the angle at which maximum power was generated during the middle phase of design. The report described it as follows:

> *To analyze the power generation and the best orientation for mounting the solar panels, one solar panel was mounted to the lid of a box with similar dimensions to a standard smoke detector. From here, the angle between the ceiling and solar panel varied from 0 (parallel to the ceiling) to 90 degrees (perpendicular to the ceiling), and an ammeter was used to read the voltage generated by the solar panel at each orientation. This procedure was conducted in both a standard-light room (~ 250 lux) and a low-light basement (~ 50 lux).* (Report 9, Middle Phase)

Experiments were also frequent in the back-end phase. Reports documented their use to: (1) verify specifications, (2) determine feasibility, (3) compare materials, (4) ensure functionality, (5) verify prototype survivability, (6) identify failure points, (7) detect design flaws, and (8) determine time and labor requirements. For example, Report 8 documented experiments to verify whether the ball joints could achieve the required rotation to ensure compliance with specifications during the back-end phase of design. The report described it as follows:

> *The test to verify Specification 2.3.4 included attaching the lego ball joints together and moving them through their range of motion, to ensure the full 180° rotation would be possible.* (Report 8, Back-End Phase)

<u>*User Testing*</u>

User testing was relatively rare in the dataset. None of the reports described user testing during the front-end or middle phases. Only three reports (Reports 3, 6, and 9) documented human-focused testing, and these activities occurred during the back-end phase to verify that the solution met design specifications. For example, Report 9 indicated user testing to confirm that the concept was easily accessible for users. The report described it as follows:

> *To verify the solution is "easy to access - users can access the detector from 3 feet above the floor", each team member used the installation piece and a 6' extendable pole to reach the detector from a seated position… Each member was able to reach the detector at ceiling height with ease. As a result, the requirement of "easy to access" was deemed verified.* (Report 9, Back-End Phase)

Although all reports documented stakeholder engagement during the front-end phase, these interactions were not classified as user tests because they did not meet the definition of a test in terms of systematicity. Instead, they took the form of conversations and consultations intended to elicit information and feedback from sponsors or stakeholders to inform requirements, specifications, and other early design decisions. As such, while they supported design decision-making, they were not structured as tests. For example, Report 1 documented consultations with sponsors to develop a list of wants and needs, which informed requirements and specifications. The report described it as follows:

> *The team determined the requirements and specifications by consulting our sponsor, [NAME]. [NAME] is the main sponsor and target user of the design and was able to provide expert insight for the team's guidance. Through a set of interviews, the team was able to determine a list of wants and needs [NAME] envisioned for her new design.* (Report 1, Front-End Phase)

Three of nine reports documented stakeholder engagement during the middle phase; however, these interactions were not classified as tests. Instead, they were intended to: (1) assess concept viability, (2) assign weights to requirements/specifications for decision matrices, and (3) check design plausibility. For example, Report 3 documented consultations with sponsors after completing a Pugh chart to confirm the viability of the highest-scoring design.The report described it as follows:

> *"After completing analysis using the Pugh chart, the overall best design was determined to be a rack and pinion with a singular drive shaft. This design scored the highest, and through talks with our sponsor, was confirmed to be viable."* (Report 3, Middle Phase)

Discussion

Despite variation in definitions of analysis, evaluation, and testing, examination of the nine reports revealed frequent use of analytical models, computational models, and experiments, particularly in the later phases of design. This pattern is consistent with prior research indicating heavy reliance on informal experimentation and analogous cases in student design work [37]. Reports also showed extensive use of prototypes to conduct experiments, aligning with prior findings on student prototyping practices [36].

User testing was notably rare across the reports. Student teams primarily engaged stakeholders in the front-end phase (e.g., sponsors) to gather information for requirements and specifications. Several factors may contribute to this pattern, including an ill-defined initial problem that shifts stakeholder engagement toward problem scoping and requirements clarification rather than concept-level analysis, evaluation, or testing [38]. Conversely, a more rigidly defined problem or difficulty interpreting stakeholder input may also limit stakeholder involvement in testing activities [38]. Additional explanations include a preference for authoritative sources (e.g., sponsors or experts) over end users, whose feedback can be more variable and difficult to synthesize [39], [40], as well as time constraints in a one-semester course. Multi-semester capstone structures [41] may provide greater opportunity for stakeholder involvement in testing. Nonetheless, user testing remained limited across all phases.

Overall, the examination of nine reports showed that student teams conducted analysis, evaluation, and testing primarily in the middle and back-end phases of the project. In contrast, none of the reports documented analysis or evaluation in the front-end phase, and only one report documented testing during this phase. This pattern is consistent with prior research showing underutilization of prototyping and testing in early design stages [36]. Together, these results reinforce prior observations that, despite the importance of testing, systematic integration across design phases remains underemphasized in student design practice [13].

The reports showed a predominant reliance on structured evaluation methods, such as Pugh charts, for concept down-selection following concept generation. While only one report explicitly attributed this choice to course guidance, the consistent use of decision matrices across all nine reports suggests a shared preference for structured decision-making tools during this phase (i.e., all student teams relied on a similar decision-matrix-based approach when selecting final concepts). This uniformity contrasts with the greater diversity observed in analysis and testing methods. This preference may reflect the ease and speed of these tools in a time-constrained, semester-long course, where resource limitations may restrict more intensive evaluation methods. However, although these matrices provided a structured mechanism for comparison, the criteria used to populate them were often only loosely supported by empirical evidence, with student teams frequently favoring highest-scoring concepts without clear linkage to systematic analysis or testing outcomes. This is consistent with prior concerns that such methods can be sensitive to subjective weighting decisions and may not adequately capture coupled design interactions [42]. Taken together, these patterns suggest that decision matrices were primarily used as aggregation mechanisms for design judgment rather than as endpoints of evidence-based evaluation. At the same time, the findings indicate a potential missed opportunity: analysis and testing activities documented elsewhere in the reports were rarely leveraged to directly inform the quantitative or qualitative inputs to these decision frameworks. This separation between evidence generation (analysis/testing) and decision synthesis (evaluation matrices) suggests a possible disconnect in how student teams integrate different design activities within concept selection. Although student teams consistently used these tools, the pattern suggests a form of procedural awareness in which they recognized the need for structured justification of design choices. This finding points to an opportunity to strengthen pedagogical support for more evidence-integrated decision-making approaches, in which analysis and testing outputs more directly inform evaluation criteria and weighting, rather than functioning as parallel or downstream activities.

Several studies have examined structured tools as supports for novice designers navigating uncertainty. Regarding analysis, the EFRT model has been proposed as an adaptive approach for early-stage robust design, particularly when design knowledge is limited or uncertain [43]; this model enables modeling of qualitative information to assess concept robustness in early design stages. Structured selection methods such as Pugh charts are similarly used to support systematic evaluation. When used pedagogically, these methods can move student teams beyond intuition by supporting (1) clearer interpretation of requirements, (2) improved understanding of design alternatives, (3) awareness of interactions among concepts, (4) justification of relative strengths and weaknesses, and (5) generation of additional concepts [42]. However, in the reports analyzed in this study, decision matrices were often applied without clear integration of evidence from analysis or testing to inform assigned scores, suggesting a more procedural than evidence-driven use of these tools. For testing, the Concept Test Design (CTD) framework provides structured support for evaluating concepts through user interaction with low-fidelity prototypes [44]. It is designed to identify errors, surface new requirements, and improve usability during early testing stages. While these tools provide structured approaches to analysis, evaluation, and testing, their depth of application in student work remains unclear. Taken together, while these tools offer structured approaches to analysis, evaluation, and testing, our findings suggest that their potential to support evidence-based decision-making may not be fully realized in practice. This highlights an opportunity to examine how student teams move from procedural use of these tools to generating evidence-based design insights.

The findings highlight a need to better support early and continuous integration of analysis, evaluation, and testing in student design projects. Encouraging student teams to implement these activities across the full design process, rather than primarily in later stages, is important because early assessment can help identify concept flaws and reduce downstream errors. However, the analysis suggests that course-specific constraints may shape how these activities are enacted, including the methods used and their timing. Targeted resources to support planning and execution of these activities may help student teams balance technical feasibility with structured decision-making.

A limitation of this study is that we did not directly interact with the student teams whose reports were being reviewed. This resulted in a limited ability to capture student teams' perspectives on analysis, evaluation, and testing modes and decision rationale or follow up to clarify ambiguous content within the reports. Moreover, the selected sample of nine reports may not be a comprehensive representative of the breadth of student teams' analysis, evaluation, and testing activities. Furthermore, we were unable to gather information regarding the student teams prior design experiences, making it difficult to determine whether decisions were influenced by previous coursework or existing skill sets. To address these limitations, future research should focus on conducting interviews with the student teams. This approach would provide richer insight on how student teams approach analysis, evaluation, and testing and the rationale behind their decisions, helping validate findings from this document analysis.

The findings highlight a clear need to better promote continuous and early analysis, evaluation, and testing activities within student design projects. Encouraging student teams to conceptualize and implement these activities across an entire design process, rather than solely in the later phases, is critical, as early assessment can help identify concept flaws and reduce downstream errors. However, course-specific constraints may dictate the way these activities are approached, including the methods chosen and the timing of their execution. Moreover, the findings suggest that providing targeted resources to help student teams plan and execute these activities would help mitigate barriers to effective analysis, evaluation, and testing and better support teams in balancing between technical feasibility and structured decision-making.

Conclusion

This qualitative document analysis of nine capstone design reports investigated student teams' analysis, evaluation, and testing activities during design work. Analysis was defined as the use of analytical or computational models to simulate concept performance under specified conditions; evaluation as the assessment of concepts against client objectives and specifications; and testing as the collection of data on how a concept performs in practice. Results showed that analysis, evaluation, and testing were infrequently documented in the front-end phase, with most activity occurring in the middle and back-end phases. In these later phases, experiments, analytical models, and computational models were most commonly used. Decision matrices such as Pugh charts were frequently used for evaluation, with student teams typically selecting concepts with the highest scores, even when the basis for these scores appeared to be primarily subjective rather than empirically grounded. These findings suggest that although student teams recognize the importance of analysis, evaluation, and testing, they may lack strategies for integrating these activities early and continuously. Because early assessment is critical for reducing downstream design risk, strengthening support for continuous engagement with these activities would benefit engineering design education.

References

[1] *Engineering Design: A Project-Based Introduction, Fourth Edition*. Accessed: Mar. 27, 2026. [Online]. Available: https://learning.oreilly.com/library/view/engineering-design-a/9781118324585/

[2] J. B. Nikander, L. A. Liikkanen, and M. Laakso, “The preference effect in design concept evaluation,” *Des. Stud.*, vol. 35, no. 5, pp. 473–499, Sep. 2014, doi: 10.1016/j.destud.2014.02.006.

[3] P. D. T. O’Connor, *Test Engineering: A Concise Guide to Cost-effective Design, Development and Manufacture*. John Wiley & Sons, 2001.

[4] Cross, Nigel, *Engineering Design Methods: Strategies for Product Design*. John Wiley & Sons, 2021.

[5] K. Tahera and C. Earl, “Testing and PLM: Connecting Process and Product Models in Product Development,” 2018. [Online]. Available: https://www.intechopen.com/chapters/63080

[6] A. Albers, J. Reinemann, T. Hirschter, J. Fahl, and N. Heitger, “Validation-Driven Design in the Early Phase of Product Development,” *Procedia CIRP*, vol. 84, pp. 630–637, Jan. 2019, doi: 10.1016/j.procir.2019.04.211.

[7] S. Ahmed and K. M. Wallace, “Identifying and supporting the knowledge needs of novice designers within the aerospace industry,” *J. Eng. Des.*, vol. 15, no. 5, pp. 475–492, Oct. 2004, doi: 10.1080/09544820410001708430.

[8] V. Viswanathan and J. Linsey, “Spanning the complexity chasm: A research approach to move from simple to complex engineering systems,” *Artif. Intell. Eng. Des. Anal. Manuf.*, vol. 28, no. 4, pp. 369–384, Nov. 2014, doi: 10.1017/S0890060414000535.

[9] V. Horvath and J. Harris, “Student experiences of interdisciplinary teamwork,” *Proc. Can. Eng. Educ. Assoc. CEEA*, Jul. 2025, doi: 10.24908/pceea.2025.19589.

[10] “Mapping problem and requirements to final solution: A document analysis of capstone design projects.” Accessed: Feb. 19, 2026. [Online]. Available: https://journals.sagepub.com/doi/epdf/10.1177/0306419018780741?src=getftr&utm_source=scopus&getft_integrator=scopus

[11] G. Pahl *et al.*, *Engineering Design: A Systematic Approach*. London, UNITED KINGDOM: Springer London, Limited, 2007. Accessed: Jan. 19, 2026. [Online]. Available: http://ebookcentral.proquest.com/lib/umichigan/detail.action?docID=972988

[12] G. Dieter and L. Schmidt, *Engineering Design*, 4th ed. 2009.

[13] K. Tahera, D. C. Wynn, C. Earl, and C. M. Eckert, “Testing in the incremental design and development of complex products,” *Res. Eng. Des.*, vol. 30, no. 2, pp. 291–316, Apr. 2019, doi: 10.1007/s00163-018-0295-6.

[14] “Knovel - 10. Verification, Validation, and Testing.” Accessed: Mar. 30, 2026. [Online]. Available: https://app.knovel.com/kn/resources/kt003Z44D1/kpHSPMAAP9/pdf?b-toc-cid=kpHSPMAAP9&b-toc-root-slug=handbook-simulation-principles&b-toc-title=Handbook%20of%20Simulation%20-%20Principles,%20Methodology,%20Advances,%20Applications,%20and%20Practice&b-toc-url-slug=part-i-principles&kpromoter=federation

[15] A. R. Patel, A. Murphy, J. D. Summers, and K. Tahera, “Testing in Engineering Design: What Are We Teaching,” *Proc. Des. Soc.*, vol. 2, pp. 2363–2372, May 2022, doi: 10.1017/pds.2022.239.

[16] G. W. Burchill, “Concept Engineering: an investigation of TIME vs. MARKET orientation in product concept development.”.
[17] C. L. Dym, *Engineering design: A Synthesis of Views*. Cambridge University Press, 1994.
[18] A. Q. Fang, “Student Team Decision-Making Processes and Assessment Criteria in Early Stage Engineering Design”.
[19] J. Alvarado, “AC 2008-1278: DESIGN, BUILD AND TEST: AN APPROACH FOR A CAPSTONE DESIGN COURSE IN ENGINEERING TECHNOLOGY”.
[20] D. Türk, B. Leutenecker, and M. Meboldt, “EXPERIENCE THE RELEVANCE OF TESTING IN ENGINEERING DESIGN EDUCATION”.
[21] A. Geras, M. Smith, and J. Miller, “A prototype empirical evaluation of test driven development,” in *10th International Symposium on Software Metrics, 2004. Proceedings.*, Chicago, IL, USA: IEEE, 2004, pp. 405–416. doi: 10.1109/METRIC.2004.1357925.
[22] “New Agile Testing Modes.” Accessed: Feb. 16, 2026. [Online]. Available: https://scialert.net/abstract/?doi=itj.2012.707.712
[23] A. Cauevic, S. Punnekkat, and D. Sundmark, “Quality of Testing in Test Driven Development,” in *2012 Eighth International Conference on the Quality of Information and Communications Technology*, Lisbon, Portugal: IEEE, Sep. 2012, pp. 266–271. doi: 10.1109/QUATIC.2012.49.
[24] J. W. Creswell, *Research design: qualitative, quantitative, and mixed methods approaches*, 4. ed. Los Angeles, Calif.: SAGE, 2014.
[25] H. Morgan, “Conducting a Qualitative Document Analysis,” *Qual. Rep.*, 2022, doi: 10.46743/2160-3715/2022.5044.
[26] “Qualitative Research Designs, Sample Size and Saturation: Is Enough Always Enough?,” *J. Advocacy Res. Educ.*, vol. 8, no. 3, Dec. 2021, doi: 10.13187/jare.2021.3.60.
[27] L. Siller, A. Roberts, M. Waier, and A. Kassouf, “Design a Better Smoke Detector,” May 2022, Accessed: Jan. 15, 2026. [Online]. Available: https://hdl.handle.net/2027.42/172471
[28] J. Neville, V. Engineering, and H. Industries, “Team 12: Heller Industries Reflow Oven Drive Train Spencer Dickhudt, Connor Hohl, Eric Kent, Nathan Kleb, Ruess Rolloff Prof. Luntz - Section 004”.
[29] A. Smith, C. Bryant, S. T. Chhetri, and D. Freeland, “Micro Capillary Tube Grinding Team 14 Winter 2023 MECHENG 450 Design Project April 24, 2023”.
[30] K. Bautista, M. Lee, J. Marshall, and M. Solar, “CoasterCraft: Prototyping for Thrills,” Dec. 2024, Accessed: Jan. 15, 2026. [Online]. Available: https://hdl.handle.net/2027.42/196259
[31] A. Ferrara, L. Hetrick, R. Lee, and M. May, “Baja SAE Adaptive Suspension,” Dec. 2024, Accessed: Jan. 15, 2026. [Online]. Available: https://hdl.handle.net/2027.42/196266
[32] J. Gottesman, B. Nordbeck, and N. Tancongco, “IMPROVING BIO-LOGGING TAG ATTACHMENT MECHANISMS FOR MARINE RESEARCH WITH BOTTLENOSE DOLPHINS”.
[33] B. Redgate, C. Olsen, A. Zhang, and A. Sarro, “Bovine Biomimicry: Making Renewable Natural Gas from Food Waste, Wastewater, and Sludge”.
[34] M. Davis, V. Dawidowicz, M. Gates, and I. Hester, “EQUIPMENT TO INVESTIGATE DOLPHIN ECHOLOCATION”.
[35] A. Alba, N. Aliff, Y. Huang, and J. Lee, “Object Detection with an Anthropomorphic Hand,” May 2025, Accessed: Jan. 15, 2026. [Online]. Available: https://hdl.handle.net/2027.42/197408

[36] M. Deininger, S. R. Daly, K. H. Sienko, and J. C. Lee, “Novice designers’ use of prototypes in engineering design,” *Des. Stud.*, vol. 51, pp. 25–65, Jul. 2017, doi: 10.1016/j.destud.2017.04.002.

[37] N. Cross, “Expertise in design: an overview,” *Des. Stud.*, vol. 25, no. 5, pp. 427–441, Sep. 2004, doi: 10.1016/j.destud.2004.06.002.

[38] I. Mohedas, S. R. Daly, R. P. Loweth, and K. H. Sienko, “Changes to stakeholder engagement approaches throughout a capstone engineering design course,” *Int. J. Technol. Des. Educ.*, vol. 34, no. 2, pp. 759–781, Apr. 2024, doi: 10.1007/s10798-023-09833-x.

[39] I. Mohedas, S. R. Daly, and K. H. Sienko, “Design Ethnography in Capstone Design: Investigating Student Use and Perceptions”.

[40] R. Wertz, M. Ross, M. Fosmire, M. Cardella, and S. Purzer, “Do Students Gather Information to Inform Design Decisions? Assessment with an Authentic Design Task in First-Year Engineering,” in *2011 ASEE Annual Conference & Exposition Proceedings*, Vancouver, BC: ASEE Conferences, Jun. 2011, p. 22.508.1-22.508.14. doi: 10.18260/1-2--17789.

[41] D. Schaefer, G. Coates, and C. Eckert, Eds., *Design Education Today: Technical Contexts, Programs and Best Practices*. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-17134-6.

[42] L. Lønmo and G. Muller, “7.1.2 Concept Selection ‐ Applying Pugh Matrices in the Subsea Processing Domain,” *INCOSE Int. Symp.*, vol. 24, no. 1, pp. 583–598, Jul. 2014, doi: 10.1002/j.2334-5837.2014.tb03169.x.

[43] J. Li, D. Horber, P. Grauberger, S. Goetz, S. Wartzack, and S. Matthiesen, “Supporting early robust design for different levels of specific design knowledge: an adaptive method for modeling with the Embodiment Function Relation and Tolerance model,” *Des. Sci.*, vol. 10, p. e46, Jan. 2024, doi: 10.1017/dsj.2024.48.

[44] C. C. Thales Martins Ponciano, “Concept test method to assist engineering students to involve users in a user-centered design project,” *Concept Test Method Assist Eng. Stud. Involve Users User-Centered Des. Proj.*, vol. 19, no. 1, pp. 0–0, Apr. 2021, doi: 10.4322/pmd.2021.002.

Acknowledgements

The author(s) received no financial support for the research, authorship, or publications of this study.

AI Use Disclosure

The authors used AI tools to assist with editing grammar, clarity, and style. All ideas, analysis, and conclusions presented in this manuscript remain the authors’ own.