Modeling and Simulation in Engineering: A Case Study of Multidisciplinary Learning

Dr. Wenbing Zhao, Cleveland State University

Wenbing Zhao received his Ph.D. in Electrical and Computer Engineering at University of California, Santa Barbara, in 2002. Dr. Zhao has a Bachelor of Science degree in Physics in 1990, and a Master of Science degree in Physics in 1993, both at Peking University, Beijing, China. Dr. Zhao joined Cleveland State University (CSU) faculty in 2004 and is currently a Professor in the Department of Electrical and Computer Engineering at CSU. Dr. Zhao published over 300 peer-reviewed papers in the area of distributed systems, smart and connected health, computer vision, machine learning, physics, and engineering education. Dr. Zhao's research is supported in part by the US National Science Foundation, the US Department of Energy, the US Department of Education. Dr. Zhao is currently serving on the organizing committee and the technical program committee for numerous international conferences. He is an Associate Editor for IEEE Access and for MDPI Computers. Dr. Zhao is a senior member of IEEE.

Dr. Yong Tao P.E., Cleveland State University

Dr. Yong Tao holds the Betty L. Gordon Endowed Chair and Distinguished Professor of the Department of Mechanical Engineering and is the past department chair at CSU. He received his Ph.D. in mechanical engineering from the University of Michigan. Prior to joining CSU in 2019, Dr. Tao has served in higher education for three decades with different roles as professor, research director, chair, associate dean, and dean in several public or private research universities. He has been active in research of the fundamentals of thermal sciences, renewable energy applications in buildings and energy system modeling and simulations. His research at CSU involves collaboration with industries on district energy systems. Dr. Tao founded Center for Integrated Modeling for Energy, Sustainability and Resiliency (IERS Center) at CSU. His recent endeavors include a Department of Education grant to establish the experiential learning environment in modeling and simulation of energy system engineering, and coordination of development of professional microcredential and certification courses in microgrid and power engineering.

Dr. Xiongyi Liu

Xiongyi Liu is an Associate Professor in the School of Education and Counseling at Cleveland State University (CSU), USA. Dr. Liu obtained her Ph.D. in Educational Psychology from University of Nebraska-Lincoln, USA. Prior to joining CSU, she has worked as research faculty at the Center for Research and Evaluation, University of Maine-Orono. Her research interests include educational technology, motivation, self-regulation, peer assessment, teacher education, STEM education, and artificial intelligence in education. At CSU, Dr. Liu is affiliated with TechHub and the Center for Integrated Modeling for Energy, Sustainability and Resiliency. She has worked on multiple interdisciplinary research projects as co-PI and evaluator for the development of innovative approaches to engineering education at various educational levels in such critical areas as modeling and simulation, cybersecurity, and AI.

Assessing Multidisciplinary Learning in a Graduate Course on Fundamental of Modeling and Simulation: A Mixed-Methods Case Study

Abstract

This paper presents a case study of multidisciplinary learning in a graduate-level modeling and simulation course. Multidisciplinary learning is understood as the integration of knowledge and methods from statistics, computer science, engineering, and domain-specific contexts to analyze and solve complex systems.

We use a mixed-methods approach combining quantitative survey measures and qualitative project analysis to examine how students develop multidisciplinary competencies. Four research questions guide the study: (1) students' perceived competence gains across key modeling and simulation domains; (2) the influence of instructional components such as peer teaching, homework, and team projects; (3) the extent to which student projects demonstrate integration of disciplinary and methodological elements; and (4) student suggestions for improving future multidisciplinary learning experiences.

Findings reveal strong perceived gains across all modeled domains, a collective (but not component-specific) association between course activities and competence, clear patterns of methodological integration centered on discrete-event simulation and related tools, and student recommendations emphasizing improved scaffolding and example-driven instruction. Together, these results offer actionable insights for strengthening multidisciplinary learning in modeling and simulation curricula.

Introduction

Modeling and simulation (M&S) play a central role in mechanical engineering by integrating mathematical modeling, numerical methods, statistical analysis, and computational tools to represent and analyze complex systems. Courses in M&S typically span discrete-event simulation, continuous modeling, Monte Carlo methods, and data-driven analysis, supported by programming and visualization. These techniques are widely applied in manufacturing, thermal–fluid systems, logistics, and mechanical system design.

Prior work highlights that simulation practice inherently blends disciplinary perspectives: effective modeling requires probability and statistics, algorithmic reasoning, physics-based system understanding, and critical evaluation of model credibility [1–6]. This integrative nature makes M&S an ideal environment for cultivating multidisciplinary competence in mechanical engineering programs.

Although multidisciplinary learning has been widely discussed in engineering education, existing studies often examine project-based learning, active learning, or conceptual development in isolation. Less is known about how students perceive their own multidisciplinary competence in an M&S context or how specific instructional elements, *e.g.,* lectures, homework, peer teaching, or semester-long projects, shape that development. Likewise, analyses of student project work rarely investigate how disciplinary and methodological elements are structurally combined.

This study addresses these gaps through a mixed-methods case analysis of a graduate-level M&S course. We integrate (i) survey-based measurement of students' perceived competence across core M&S domains, (ii) regression analysis of how instructional activities relate to those perceptions, and (iii) structural analysis of student projects to characterize how methods and disciplinary perspectives are integrated in practice. The goal is to provide an empirically grounded account of how M&S coursework supports multidisciplinary learning and to offer insights for course and curriculum design.

Four research questions guide the study:

- RQ1: How do students perceive the development of multidisciplinary competencies through modeling and simulation courses?
- RQ2: How do course components—such as peer teaching, team projects, and homework—relate to perceived multidisciplinary learning?
- RQ3: In what ways do student projects reflect integration of knowledge and methods from multiple disciplines?
- RQ4: What suggestions do students offer to improve multidisciplinary learning experiences?

RQ1 examines self-reported gains across modeling and simulation domains. RQ2 investigates how instructional elements support these gains. RQ3 evaluates the disciplinary and methodological integration evident in student projects. RQ4 captures student recommendations for strengthening future offerings. Together, these questions offer a comprehensive view of multidisciplinary learning within a modeling and simulation course.

1 Methods

This study was conducted in a graduate-level mechanical engineering course on modeling and simulation. The course combines statistical analysis, computational modeling, and simulation paradigms through lectures, programming-based homework, peer teaching, and a semester-long project. A mixed-methods design was used to examine how students developed multidisciplinary competence across these experiences.

1.1 Participants and Instructional Setting

Thirty-five students enrolled in the course, forming ten teams to complete semester-long modeling and simulation projects. Project topics spanned traffic operations, Computer Numerical Control (CNC) machining, supermarket checkout systems, queueing and logistics models, drone delivery and vertiport operations, electric vehicle modeling, baseball trajectory analysis, and manufacturing defect prediction. These projects required students to construct a full modeling workflow, including system definition, model implementation, experimentation, and analysis.

The course ran over a full academic semester. Students engaged continuously with modeling tools through homework, peer teaching, in-class activities, and project development. A single post-course survey was administered; no pre-assessment was used because our focus was on students' retrospective perceptions after completing the full modeling experience.

1.2 Data Sources

Three types of data were collected through the post-course survey (30 valid responses) and student project descriptions:

- The Course Activities Index (CAI) is based on student ratings of instructional components, including lectures, homework, peer teaching, and the team project, on a 1–5 usefulness scale. These ratings provide a process-oriented measure of engagement across key course elements.
- The Multidisciplinary Competence Index (MCI) captures students' perceived improvement across eight modeling and simulation domains: vocabulary, foundational statistics, input/output analysis, discrete-event simulation, continuous simulation, Monte Carlo methods, visualization, and verification/validation. Items were rated on a 1–5 improvement scale, with an NA option when applicable. The MCI score is the mean of the eight items.
- Open-Ended Responses. Students described which aspects of the course helped them most or least and provided suggestions for improvement. These responses offer qualitative insight beyond numerical scales.
- Project Descriptions. Team-submitted project titles and abstracts were used to analyze disciplinary breadth and methodological integration across student work.

1.3 Analytic Procedures

For RQ1, descriptive statistics were computed for the MCI, and reliability was assessed using Cronbach's alpha [7] and corrected item–total correlations.

For RQ2, we estimated ordinary least squares regressions predicting MCI from the four CAI components. Heteroskedasticity-consistent (HC3) [8, 9] standard errors were used. Influence diagnostics (externally studentized residuals, leverage, and Cook's distance) identified observations with disproportionate influence; results were compared across full and cleaned samples.

For RQ3, project descriptions were tagged using controlled vocabularies for disciplines and modeling concepts. Frequency summaries were followed by attempts at k-means and spectral clustering. Because clusters lacked meaningful separation, we used network-based visualization (three-band maps, heatmaps, chord diagrams, and Sankey diagrams) to characterize relationships among disciplines, concepts, and projects.

For RQ4, open-ended responses were analyzed using a short-text natural language processing (NLP) pipeline [10]. Responses were preprocessed, embedded with a sentence-transformer model, reduced via uniform manifold approximation and project (UMAP) [11], and clustered with

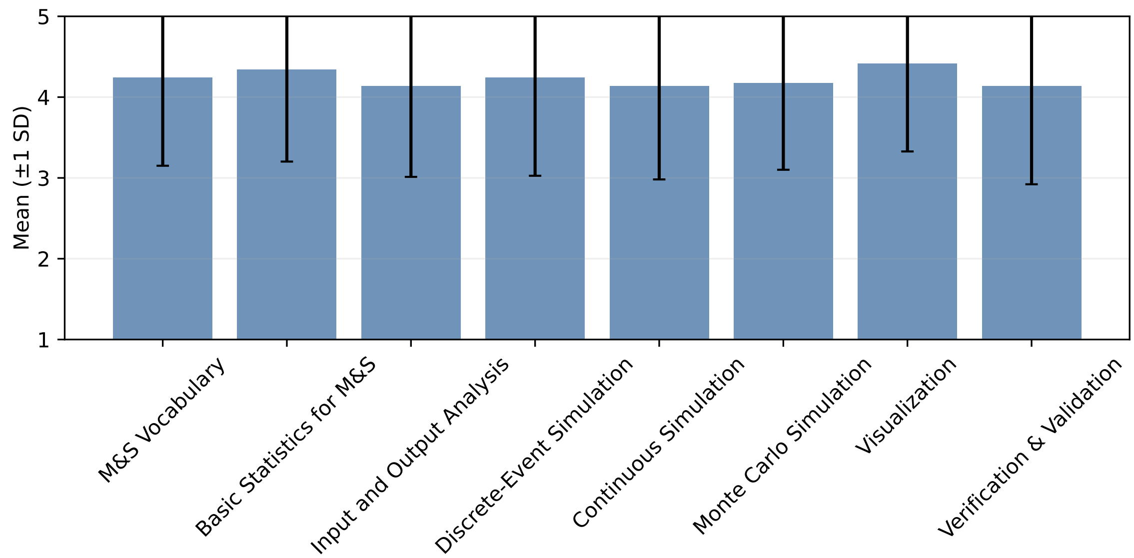

Figure 1: Mean ratings and standard deviations for the eight modeled domains.

hierarchical density-based spatial clustering of applications with noise (HDBSCAN) [12]. Cluster labels were generated using class-based TF-IDF (short for term frequency and inverse document frequency) [13] and keyphrase extraction and then verified through manual review.

For clarity, clusters represent semantically cohesive sets of responses in the embedding space, not equal mention counts of particular terms. Cluster labels are composite descriptors formed from class-specific term salience (class-based TF-IDF and keyphrase extraction) and manual validation; they indicate the dominant, co-mentioned ideas within each theme.

RQ1: How do students perceive the development of multidisciplinary competencies through modeling and simulation courses?

Perceived competence development was measured using the Multidisciplinary Competence Index (MCI), derived from eight Likert-scale [14] items evaluating improvement across core modeling and simulation domains: M&S vocabulary, basic statistics, input–output data analysis, discrete-event simulation, continuous simulation, Monte Carlo methods, visualization, and verification and validation. This composite provides a concise, student-reported indicator of multidisciplinary learning.

Figure 1 displays the item means with standard deviations. All domains show strong perceived gains, with means above 4.0 on the 1–5 scale. Foundational skills such as visualization and basic statistics are rated highest, while more advanced topics (*e.g.,* Monte Carlo simulation, verification & validation) exhibit similarly positive ratings. The narrow variability across items suggests consistent improvement across students.

Figure 2 shows the overall MCI distribution. The histogram and density estimate indicate a top-heavy pattern concentrated in the 4–5 range, consistent with strong perceived competence. A small number of outliers are visible using Tukey's IQR (short for interquartile range) rule; these represent students who reported notably lower perceived gains.

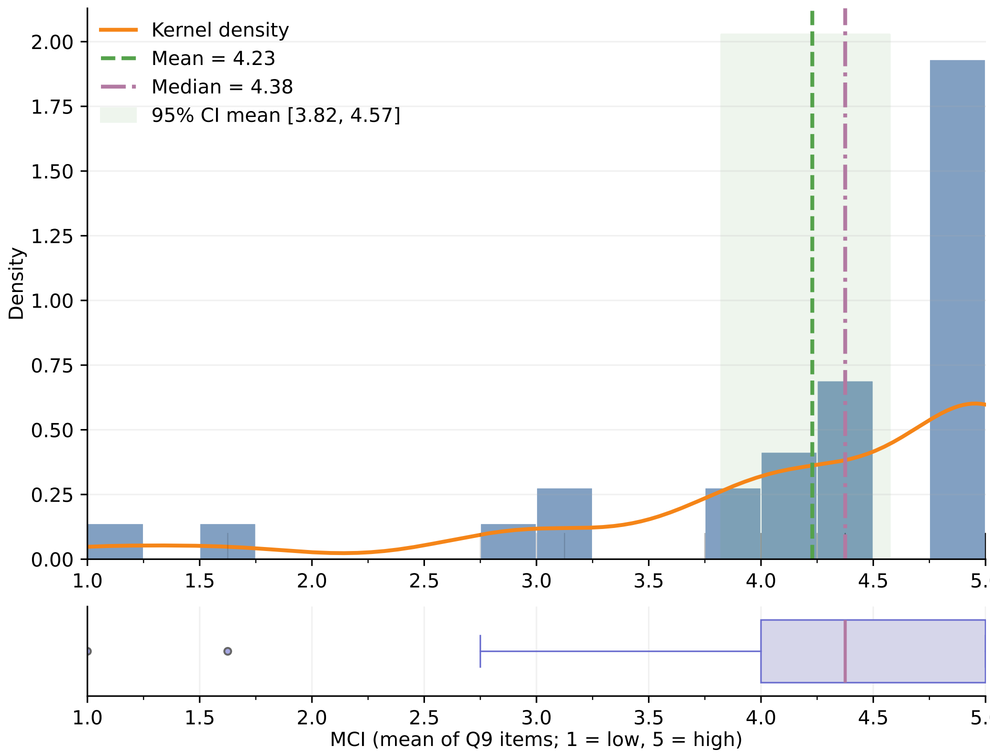

Figure 2: Distribution of the MCI with histogram, kernel density, and boxplot.

To evaluate the reliability of the MCI, we computed Cronbach's α and corrected item–total correlations (Table 1). Cronbach's α of 0.976 indicates excellent internal consistency, and all item–total correlations exceed 0.85. These results confirm that the eight domains function cohesively as a measure of integrated competence development.

To contextualize these self-reported gains, we note that students engaged in ten semester-long modeling and simulation projects spanning diverse domains such as traffic systems, CNC machining, supermarket checkout modeling, multi-server queueing systems, infantry logistics, drone operations, electric vehicles, baseball trajectory analysis, and manufacturing defect prediction. These projects required students to construct complete modeling workflows, apply analytical and computational tools, and interpret simulation outputs. Because the domains represented in the projects align closely with the MCI skill areas, the strong MCI results can be interpreted as reflecting students' engagement in authentic, multidisciplinary modeling practice.

RQ2: How do course components, such as peer teaching, team projects, and homework, impact multidisciplinary learning?

To examine how course activities relate to perceived competence, we use the Course Activities Index (CAI), which consolidates student ratings of lectures, homework, peer teaching, and the team project on a 1–5 scale. CAI represents students' engagement with key instructional elements, complementing the MCI as an outcome measure. This approach aligns with prior research showing that peer instruction, project-based learning, and iterative practice can influence conceptual understanding and engagement in engineering courses [15–18].

Table 1: Reliability summary for the eight MCI domains.

Item	Mean	SD	Item_Total_Corr
M&S Vocabulary	4.240	1.090	0.912
Basic statistics for M&S	4.340	1.140	0.931
Input/output data analysis	4.140	1.130	0.920
Discrete-event simulation	4.240	1.210	0.912
Continuous system simulation	4.140	1.160	0.853
Monte Carlo simulation	4.170	1.070	0.922
Visualization fundamentals	4.410	1.090	0.901
Verification and validation	4.140	1.220	0.876

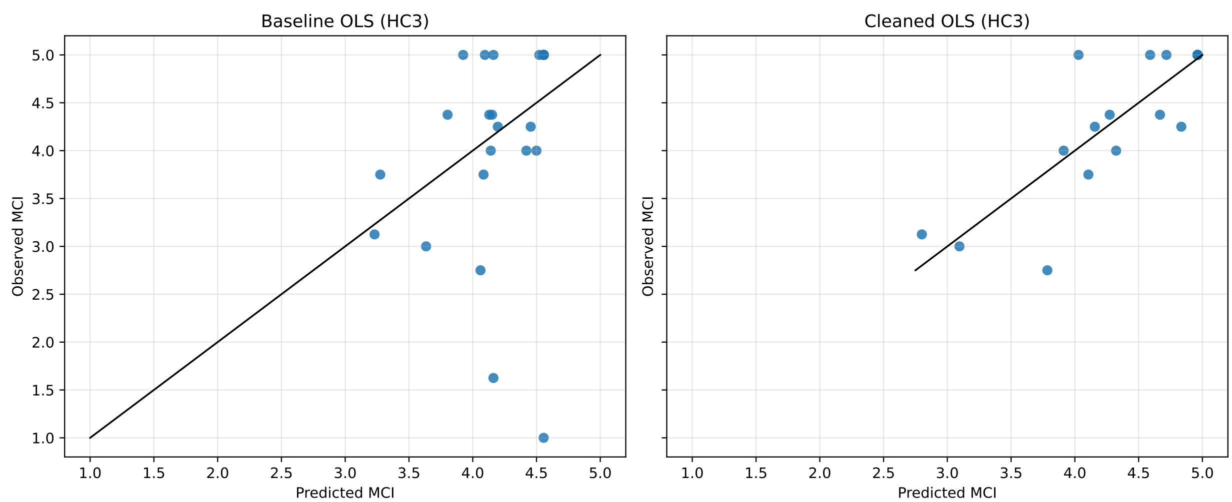

Figure 3: Observed vs. predicted MCI for baseline OLS (left) and cleaned OLS (right). The $y = x$ line marks ideal calibration.

We estimated an ordinary least squares (OLS) regression with MCI as the dependent variable and all four CAI components as predictors. Because cross-sectional survey data often exhibit heteroskedasticity, we report heteroskedasticity-consistent (HC3) standard errors to ensure valid inference in a modest sample [9, 19]. To assess model robustness, we applied standard influence diagnostics (externally studentized residuals, leverage, and Cook's D) and compared the full-sample results with a "cleaned" model that excludes influential observations.

Figure 3 contrasts the baseline and cleaned fits. The baseline model shows wide dispersion around the identity line, whereas the cleaned fit exhibits substantially tighter alignment, indicating improved calibration once influential points are removed.

Figure 4 shows the 95% confidence intervals (CI) for each CAI component under both models. Across all fits, all intervals cross zero, indicating that no individual component—lecture, homework, peer teaching, or project—exhibits a statistically reliable partial association with MCI when the others are held constant. This pattern likely reflects (1) correlations among CAI components (*i.e.,* multicollinearity) and (2) limited sample size, both of which inflate standard

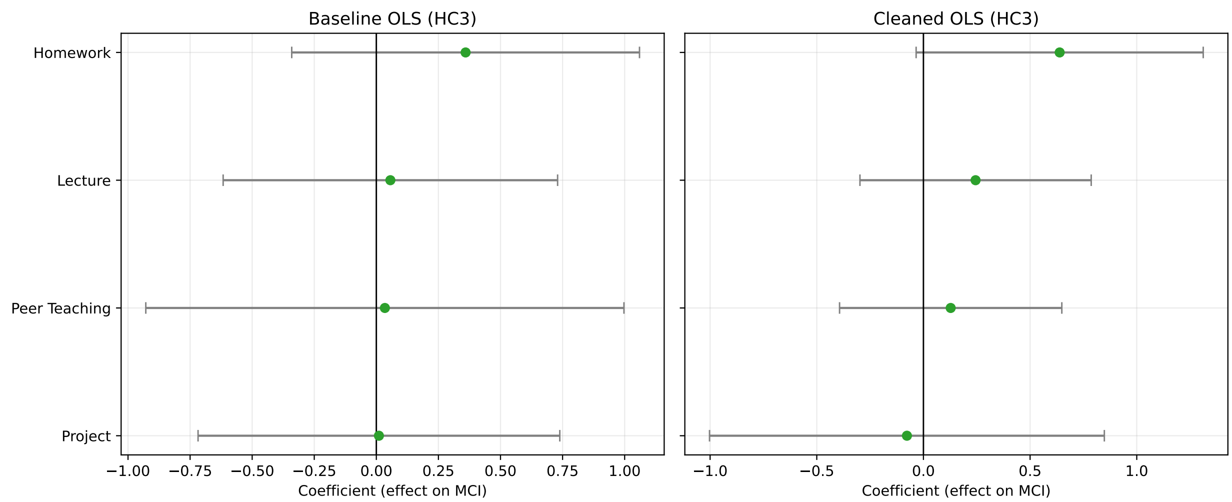

Figure 4: Coefficients with 95% CIs for baseline OLS (left) and cleaned OLS (right).

Table 2: Regression metrics for predicting MCI from CAI components.

Model	n	p	R^2	Adj R^2	RMSE	MAE	F-p
OLS (HC3)	30	4	0.131	-0.008	0.958	0.650	0.165
Cleaned OLS (HC3)	24	4	0.742	0.688	0.358	0.224	0.000

errors and reduce precision.

Table 2 summarizes overall model performance. The baseline model shows minimal explanatory power ($R^2 = 0.131$; adjusted $R^2 < 0$) and poor calibration. After excluding six influential points, the cleaned model demonstrates substantially stronger fit ($R^2 = 0.742$, adjusted $R^2 = 0.688$) and lower error (RMSE $= 0.358$). The cleaned model's joint F-test is highly significant, indicating that the four CAI components collectively explain a meaningful share of variance in perceived competence, even though no single component dominates.

Taken together, these results suggest that while students perceive the overall combination of course activities as supporting their competence development, the four individual components do not exhibit distinct partial effects when modeled simultaneously. The strong joint association in the cleaned model indicates that the instructional design, taken as a whole, relates meaningfully to perceived multidisciplinary learning, but the modest sample and correlated predictors limit the precision of activity-specific estimates.

RQ3: Integration of Knowledge and Methods in Student Projects

To examine how the student projects integrated disciplinary knowledge and modeling and simulation (M&S) methods, we analyzed each project's disciplinary tags and methodological concepts using controlled vocabularies from simulation literature [20, 21]. This enabled construction of frequency tables (Table 3) to identify commonly used disciplines and core analytical techniques.

Descriptively, all projects draw on a common simulation backbone, with Discrete-Event

Table 3: Frequency of disciplines and M&S concepts across projects.

Discipline	Count
Modeling & Simulation	10
Programming & Software	9
Logistics & Operations	6
Manufacturing & CNC	4
Physics & Engineering	3
Energy & Sustainability	2
Statistics & Probability	2

(a) Disciplines

Concept	Count
Discrete-Event Simulation	8
Queueing Theory	5
Monte Carlo Simulation	3
Optimization	3
Bayesian Inference	2
Continuous Simulation	2
Genetic Algorithms	1
Input/Output Analysis	1
Sensitivity Analysis	1

(b) M&S Concepts

Simulation and Programming & Software appearing most frequently. However, frequency alone does not reveal how disciplines and methods connect within projects.

Initial clustering attempts using project–feature vectors (disciplines + concepts) produced trivial solutions: k-means and spectral clustering both grouped nine of the ten projects into a single dominant cluster. This reflects high methodological overlap across projects and indicates that clustering does not meaningfully differentiate integration patterns.

Given the lack of separation, we used network-based visualization to examine relational structure. Figure 5 shows a three-layer map linking disciplines (inner band), M&S concepts (middle band), and projects (outer band). Discrete-Event Simulation serves as a central hub, connecting to many disciplines and nearly all projects, while concepts such as Sensitivity Analysis and Bayesian Inference appear at the periphery.

To further examine methodological structure, we constructed concept–concept co-occurrence matrices. The heatmap in Figure 6 highlights a clear methodological core: Discrete-Event Simulation and Queueing Theory co-occur in four projects, forming the strongest pairing. Other co-occurrences are sparse (mostly 0–1), indicating that additional techniques were used selectively and typically layered onto this central operational modeling framework.

The integration patterns suggested by the heatmap are reinforced by the broader structural view of method usage across projects. Discrete-Event Simulation and Queueing Theory consistently emerge as the most interconnected methods, appearing together in multiple projects and linking to a wide range of disciplinary inputs. In contrast, techniques such as Optimization, Genetic Algorithms, and Bayesian Inference appear only in isolated contexts, reflecting their roles as specialized analytical extensions rather than core modeling tools. Overall, the project ecosystem exhibits a tightly connected methodological center surrounded by a set of peripheral, low-frequency techniques.

A complementary perspective arises when examining how disciplinary backgrounds connect to modeling and simulation methods. Core disciplines, including Modeling & Simulation, Programming & Software, and Statistics & Probability, feed into a wide range of analytical

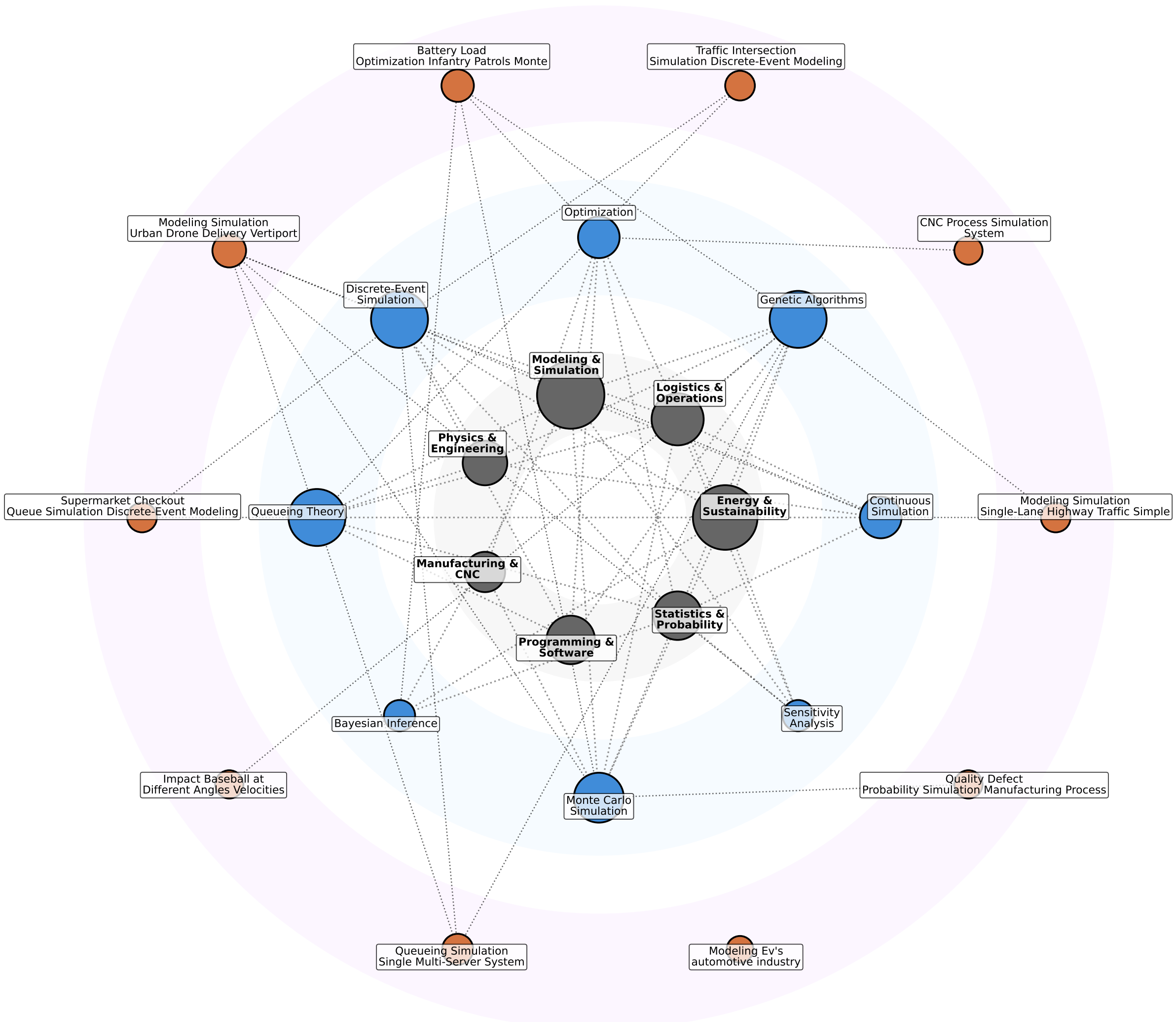

Figure 5: Three-band integration map linking disciplines, M&S concepts, and projects.

techniques, with especially strong contributions to Discrete-Event Simulation, Queueing Theory, and, to a lesser extent, Monte Carlo and Continuous Simulation. More specialized methods such as Bayesian Inference and Sensitivity Analysis draw only limited disciplinary support, underscoring their peripheral role in the overall methodological landscape. These patterns highlight how a small set of foundational disciplines underpins the majority of analytical activity across projects.

Overall, RQ3 reveals that student projects exhibit a clear integration pattern centered on operational modeling methods, particularly Discrete-Event Simulation and Queueing Theory, supported by consistent contributions from modeling, programming, and statistical disciplines. More advanced or specialized techniques (*e.g.,* Bayesian inference, sensitivity analysis, optimization) appear, but remain peripheral. The network-level analyses thus demonstrate that students blend methods and disciplinary knowledge in coherent ways, though largely around a shared methodological core.

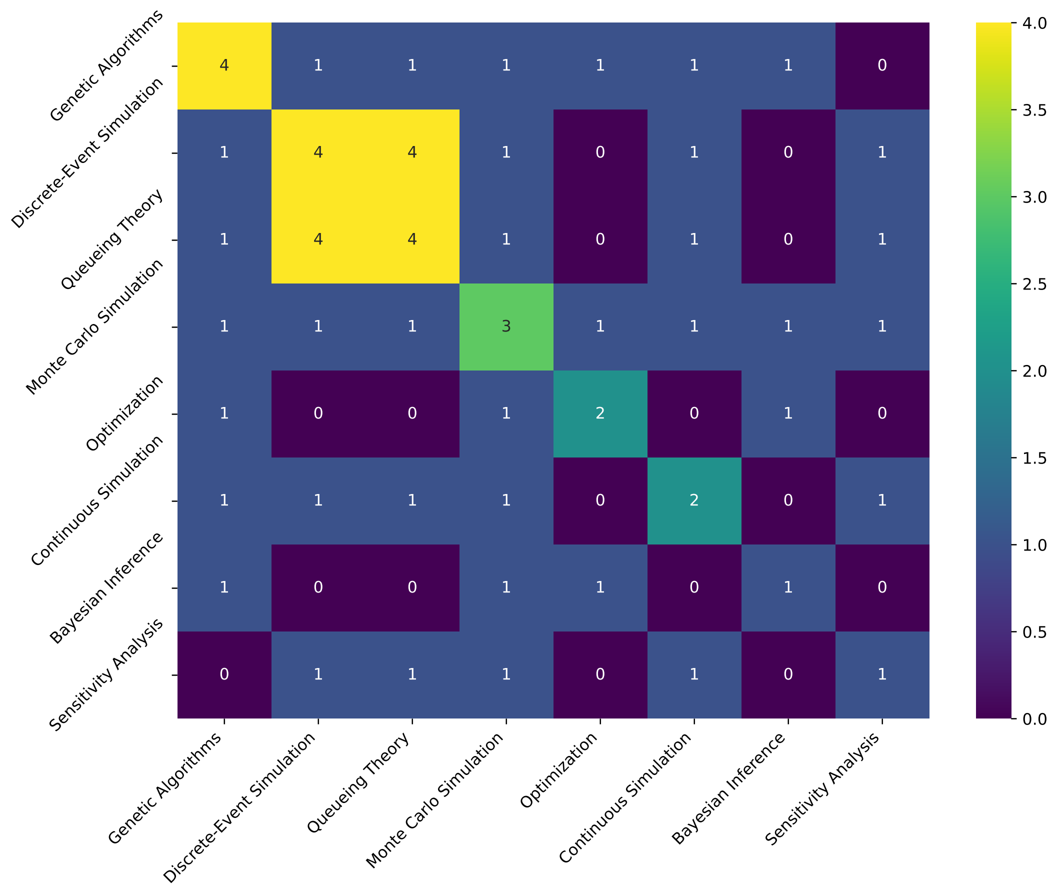

Figure 6: Concept–concept co-occurrence heatmap.

RQ4: What suggestions do students offer to improve the multidisciplinary learning experiences?

To address this research question, we analyzed student responses to two open-ended survey items: (1) aspects of the course that helped most or least with learning, and (2) suggestions for improvement. Responses were clustered using the NLP workflow described in the Methods section (preprocessing, transformer-based embeddings, UMAP reduction, HDBSCAN clustering, and cluster labeling via class-based TF-IDF and keyphrase extraction).

Aspects Students Found Most Helpful. Four coherent themes emerged from the "helped most" responses, which are illustrated at the top of portion of Figure 7. The largest theme centered on *hands-on simulation practice*, with students highlighting the value of realistic modeling tasks, simulation paradigms, and opportunities to validate models. A second theme focused on *Python-based analysis*: many students noted substantial gains in programming confidence and the ability to use Python for simulation and data analysis. The third theme combined *homework, projects, and peer teaching*, reflecting that these activities were frequently co-mentioned as mutually reinforcing components that required applying concepts in practice. The final theme involved *early-course exposure to Python and simulation fundamentals*, which students credited for building initial confidence and supporting later project work.

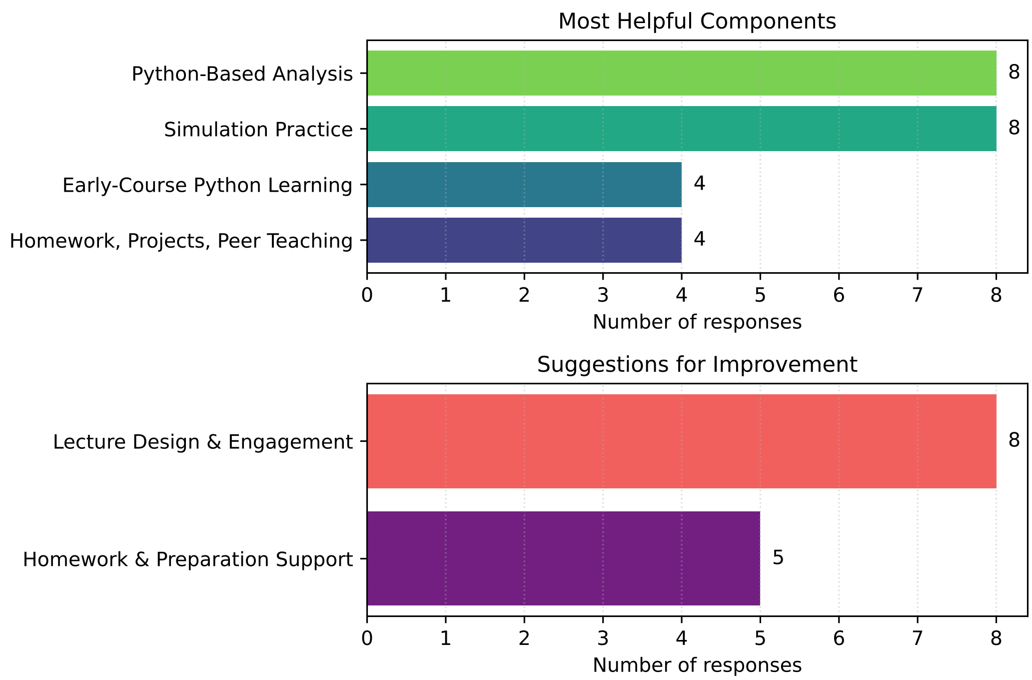

Figure 7: Themes of most helpful course components.

Aspects Students Found Least Helpful. The "least helpful" responses did not form any stable clusters. HDBSCAN labeled all responses as noise due to their sparsity and heterogeneity. Manual inspection confirmed that these comments were highly individualized (e.g., requests for more examples in specific lectures) and did not represent shared concerns. The absence of coherent negative themes suggests that no particular component was consistently viewed as ineffective.

Suggestions for Course Improvement. Analysis of improvement suggestions yielded two meaningful themes, as shown at the bottom portion of Figure 7). The first, *Homework & Preparation Support*, included requests for earlier release of assignments, additional programming tutorials, and more structured guidance during technical tasks. The second, *Lecture Design & Engagement*, emphasized interest in more interactive coding demonstrations, greater in-class practice time, and clearer alignment between lectures and hands-on activities. A third cluster was algorithmically identified but consisted almost entirely of non-suggestions (*e.g.,* "no suggestions," "everything is great") and was excluded from interpretation.

Synthesis. Overall, students valued experiential modeling work, iterative assignments, and opportunities to build computational fluency. Their suggestions focused not on changing the course's conceptual structure, but on strengthening scaffolding and instructional support, particularly for programming and example-driven explanations. These recommendations align with established principles in simulation pedagogy, where learning is enhanced when conceptual content, computational practice, and real-world application are closely integrated.

Conclusive Remarks and Future Work

This study provides an integrated, mixed-methods account of how a graduate-level modeling and simulation course supports multidisciplinary learning. Quantitative analyses of perceived competence were complemented by regression-based examination of course activities, while structural and visual analyses of student projects illuminated how methods and disciplines connect in practice. Open-ended reflections, analyzed through a transparent text-analytics pipeline, added further insight into students' learning experiences. Together, these approaches show (1) strong perceived gains across core modeling and simulation domains, (2) conditions under which activity-level regression inferences are credible, (3) clear network-level patterns in how students integrate modeling techniques, and (4) actionable recommendations for course design and instructional support.

For RQ1, students reported high perceived improvement across all eight modeled domains, with item means above 4.0 and excellent internal consistency. These results indicate that students experienced competence growth as an integrated whole rather than as isolated skills.

For RQ2, baseline regression results showed weak and inconclusive associations between course components and perceived competence. However, after removing a small number of influential observations, the cleaned model exhibited substantially improved fit, suggesting that the instructional components jointly contribute to competence but do not show distinct partial effects when modeled simultaneously. This finding highlights the need for diagnostic checks and robust inference in small-sample educational analyses.

For RQ3, frequency- and feature-based clustering approaches revealed little structural separation among projects, reflecting substantial methodological overlap. Network-based visualization, however, yielded meaningful insights: a pronounced methodological core centered on discrete-event simulation and queueing theory, supported by programming, modeling, and statistical disciplines. Peripheral techniques such as optimization, Bayesian inference, and sensitivity analysis appeared only in specialized contexts. These visualizations offer a richer representation of integrative patterns than clustering alone.

For RQ4, students identified four main contributors to their learning, *i.e.,* hands-on simulation practice, Python-based analysis, homework/projects/peer teaching, and early-course foundational support, alongside two targeted areas for improvement: additional scaffolding for homework and enhanced lecture design and engagement. The absence of a cohesive "least helpful" theme suggests that negative feedback was individualized rather than systemic.

Beyond this single course, the findings have broader implications for mechanical engineering curricula. Because modeling and simulation inherently draw on statistics, computation, physics-based reasoning, and domain knowledge, they represent a natural platform for cultivating multidisciplinary competence. The mixed-methods framework used here, *i.e.,* combining competence measures, activity-level modeling, and structured project analysis, offers a replicable model for departments seeking a richer understanding of how students integrate disciplinary knowledge. Instructional practices emphasized by students, including programming-based homework, iterative project work, and peer teaching, may strengthen integrative learning when adopted across other mechanical engineering contexts.

Limitations. This study is based on a modest sample from a single course offering, which limits statistical power and generalizability. The competence index, though reliable, is self-reported rather than performance-based. Coding of project descriptions and interpretation of open-ended responses involve human judgment, despite the use of controlled vocabularies and transparent procedures. Clustering limitations in RQ3 and the "non-suggestion" grouping in RQ4 highlight that density-based algorithms capture semantic proximity rather than substantive quality. Finally, visualizations describe relational structure but do not support causal claims.

Future Research. Five directions emerge: (1) collecting larger, multi-cohort datasets across institutions to improve generalizability; (2) incorporating performance-based assessments (e.g., scored models, code reviews) to triangulate self-reported gains; (3) using stronger explanatory designs—such as quasi-experimental manipulations of instructional components—to move beyond correlational inference; (4) advancing network modeling of project integration, including weighted or multiplex networks and temporal analyses; and (5) refining text analytics through hybrid transformer–rule-based pipelines and human-in-the-loop clustering to ensure themes reflect actionable content. These steps would deepen understanding of how modeling and simulation courses cultivate multidisciplinary capability and how instructional design can further enhance student learning.

References

[1] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*. McGraw-Hill, 1991, fundamentals of discrete-event simulation, input/output analysis, validation. [Online]. Available: https://books.google.com/books/about/Simulation_Modeling_and_Analysis.html?id=jeFQAAAAMAAJ

[2] G. S. Fishman, *Discrete-Event Simulation: Modeling, Programming, and Analysis*, ser. Springer Series in OR and Financial Engineering. Springer, 2001. [Online]. Available: https://link.springer.com/book/10.1007/978-1-4757-3552-9

[3] J. W. Forrester, *Industrial Dynamics*. MIT Press, 1961. [Online]. Available: https://systemdynamics.org/what-is-system-dynamics-old/

[4] F. E. Cellier and E. Kofman, *Continuous System Simulation*. Springer, 2006. [Online]. Available: https://ww2.mathworks.cn/academia/books/continuous-system-simulation-cellier.html

[5] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Proceedings of the IEEE Symposium on Visual Languages*, 1996. [Online]. Available: https://www.cs.umd.edu/~ben/papers/Shneiderman1996eyes.pdf

[6] S. M. Youngblood, D. K. Pace, P. L. Eirich, D. M. Gregg, and J. E. Coolahan, "Simulation verification, validation, and accreditation," *Johns Hopkins APL Technical Digest*, vol. 21, no. 3, 2000. [Online]. Available: https://www.jhuapl.edu/Content/techdigest/pdf/V21-N03/21-03-Youngblood.pdf

[7] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951.

[8] J. G. MacKinnon and H. White, "Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties," *Journal of Econometrics*, vol. 29, pp. 305–325, 1985.

[9] J. S. Long and L. H. Ervin, "Using heteroscedasticity consistent standard errors in the linear regression model," *The American Statistician*, vol. 54, no. 3, pp. 217–224, 2000. [Online]. Available: https://www.jstor.org/stable/2685594

[10] J. Ma and Y. Guo, "Using transformer embeddings for short-response clustering in stem education," *Journal of Educational Data Mining*, vol. 15, no. 1, pp. 1–26, 2023.

[11] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2020.

[12] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on hierarchical density estimates," *Advances in Knowledge Discovery and Data Mining*, pp. 160–172, 2015.

[13] S. Sia, S. Dalmia, and S. Mielke, "Tired of topic models? try clustering your documents with tf-idf and c-tf-idf," *Proceedings of the Workshop on Natural Language Processing for Digital Humanities*, pp. 184–196, 2020.

[14] R. Likert, "A technique for the measurement of attitudes," Ph.D. dissertation, Columbia University, 1932. [Online]. Available: https://legacy.voteview.com/pdf/Likert_1932.pdf

[15] C. H. Crouch and E. Mazur, "Peer instruction: Ten years of experience and results," *American Journal of Physics*, vol. 69, no. 9, pp. 970–977, 2001.

[16] L. Helle, P. Tynjälä, and E. Olkinuora, "Project-based learning in post-secondary education – theory, practice and rubber sling shots," *Higher Education*, vol. 51, pp. 287–314, 2006.

[17] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth, "Active learning increases student performance in science, engineering, and mathematics," *Proceedings of the National Academy of Sciences*, vol. 111, no. 23, pp. 8410–8415, 2014.

[18] J. Zilvinskis, L. M. Rocconi, and A. D. Dumford, "The relationship between student engagement and academic performance: Evidence from nsse," *Journal of College Student Development*, vol. 58, no. 6, pp. 863–879, 2017.

[19] H. White, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica*, vol. 48, no. 4, pp. 817–838, 1980.

[20] J. Banks, J. S. Carson, B. L. Nelson, and D. M. Nicol, *Discrete-Event System Simulation*, 5th ed. Prentice Hall, 2010.

[21] A. M. Law, *Simulation Modeling and Analysis*, 5th ed. McGraw-Hill Education, 2015.