Board 201: A New Public Dataset for Exploring Engineering Longitudinal Development by Leveraging Curricular Analytics

Dr. David Reeping, University of Cincinnati

Dr. David Reeping is an Assistant Professor in the Department of Engineering and Computing Education at the University of Cincinnati. He earned his Ph.D. in Engineering Education from Virginia Tech and was a National Science Foundation Graduate Research Fellow. He received his B.S. in Engineering Education with a Mathematics minor from Ohio Northern University. His main research interests include transfer student information asymmetries, threshold concepts, curricular complexity, and advancing quantitative and fully integrated mixed methods.

Dr. Matthew W. Ohland, Purdue University, West Lafayette

Matthew W. Ohland is the Dale and Suzi Gallagher Professor and Associate Head of Engineering Education at Purdue University. He has degrees from Swarthmore College, Rensselaer Polytechnic Institute, and the University of Florida. His research on the longitudinal study of engineering students and forming and managing teams has been supported by the National Science Foundation and the Sloan Foundation and his team received for the best paper published in the Journal of Engineering Education in 2008, 2011, and 2019 and from the IEEE Transactions on Education in 2011 and 2015. Dr. Ohland is an ABET Program Evaluator for ASEE. He was the 2002–2006 President of Tau Beta Pi and is a Fellow of the ASEE, IEEE, and AAAS.

Dr. Kenneth Reid, University of Indianapolis

Kenneth Reid is the Associate Dean and Director of Engineering at the R. B. Annis School of Engineering at the University of Indianapolis. He and his coauthors were awarded the Wickenden award (Journal of Engineering Education, 2014) and Best Paper award, Educational Research and Methods Division (ASEE, 2014). He was awarded an IEEE-USA Professional Achievement Award (2013) for designing the B.S. degree in Engineering Education. He is a co-PI on the ”Engineering for Us All” (e4usa) project to develop a high school engineering course ”for all”. He is active in engineering within K-12, (Technology Student Association Board of Directors) and has written multiple texts in Engineering, Mathematics and Digital Electronics. He earned a PhD in Engineering Education from Purdue University, is a Senior Member of IEEE, on the Board of Governors of the IEEE Education Society, and a Member of Tau Beta Pi.

Dr. Hossein EbrahimiNejad, Drexel University

Hossein EbrahimiNejad is a data scientist currently working with the office of Enrollment Analytics at Drexel University. He received his PhD in Engineering Education from Purdue University, where he gained a strong knowledge of higher education and strategic enrollment management. Hossein's skills in data management, data visualization, and predictive modeling allow him to empower stakeholders to make strategic decisions using advanced analytic information. With Drexel's commitment to promoting and supporting student success, Hossein's work with Enrollment Analytics allows the University to make data-driven strategic decisions regarding enrollment and financial projections that are as effective and efficient as possible.

Nahal Rashedi

A New Public Dataset for Exploring Engineering Longitudinal Development Leveraging Curricular Analytics (Year One)

Abstract

Considering the increasing demand for engineering graduates, understanding what limits students from completing their degrees has been a consistent question in the literature. The nontrivial variance in the pathways students take in obtaining an engineering degree, especially in cases where students abandon the study of engineering, suggests that longitudinal datasets can hold a wealth of information to uncover factors contributing to attrition. Accordingly, this project uses historical data to explore curricular factors that create barriers for different students by leveraging a new framework, Curricular Analytics, for quantifying the impact of such factors. This paper provides an overview of the year one activities and achievements for the NSF project, Studying Undergraduate Curricular Complexity for Engineering Student Success (SUCCESS). These activities included data collection, data verification, and drafting an R package to calculate the curricular complexity metrics.

Introduction

Studying pathways into and out of engineering is a classic area of research in engineering education, with studies highlighting various factors that influence a student's decision to pursue the associated major, stay and complete the degree, or leave; these include personal, social, and institutional factors [1]–[3]. With over one million student records, MIDFIELD has been a wealth of information for researchers interested in studying these engineering curricular pathways [4]. It is composed of four tables that are linked together using a unique identifier for each student: (1) course information, including credit hours and grades; (2) term information that describes the program and students' academic standing; (3) student demographic information; and (4) degree information for graduates such as the program and term completed [5]. Perhaps the most underexplored dataset within MIDFIELD is the course table. The course table contains approximately 52 million observations, each a course taken by a student at some point during their degree from 21 institutions – some dating back in the 1980s and as current as 2021. This course-taking data includes the grade the student earned and, for some institutions, information about the rank of the instructor and course modality.

This project combines the course-taking data from MIDFIELD with an emerging framework called *Curricular Analytics* [6]. We were motivated to synergize these ideas to quantitatively explore different students' course-taking trajectories and the curricula they engage with across institutional and disciplinary contexts. In simulated and empirical data, the metrics within Curricular Analytics negatively correlate with completion rates [7], [8], suggesting that the framework can be useful for retention research and as a curricular design tool. Thus, this project can enable a new formal strand of research to emerge from the pool of work related to curricular design. We will first briefly review the premise of Curricular Analytics.

Curricular Analytics

Curricular Analytics first appeared in its most recognizable form in [9] but was formalized in work by Heileman and colleagues [6]. The premise of the framework is to characterize the complexity of curriculum in terms of two overarching constructs, instructional complexity and structural complexity. Instructional complexity captures the latent dimensions of what makes a curriculum complex and affects student progress, which includes instructor quality, student support systems, and modality. The authors admit these are difficult to measure, especially at the curriculum level; therefore, the instructional complexity of a course uses pass rates as a proxy for the more difficult-to-measure dimensions. Although pass rates are sufficient for conducting simulation analyses of completion rates, this aspect of Curricular Complexity is generally too underdeveloped to be used as a theoretical framework on its own – despite some work by Waller [10], who used the concept of grade anomaly instead of course pass-rate as a proxy for instructional complexity.

Much more promising, however, is the idea of *structural complexity*. This construct has the researcher examine the curriculum itself by using network analysis to measure sequencing and interconnectedness in a plan of study. Two intrinsic measurements are associated with each course in Curricular Analytics: (1) the blocking factor, which counts how many courses are inaccessible to a student upon failing a specific course, and (2) the delay factor, the longest prerequisite chain through the course. When these two factors are summed, we obtain the cruciality of a course. The cruciality provides a measure of how central a course is to the overall curriculum, which translates to the ability to identify potential bottlenecks in the form of gateway courses – a form of local analysis. The sum of these factors for all courses in a plan of study is the structural complexity, which provides a global metric for the plan of study that enables comparative and correlational analyses.

Curricular Analytics is a relatively new framework with the potential to address existing research questions and generate new ones within educational research. To date, structural complexity has been used to correlate program quality with a curriculum's structural complexity [11] and predict four-, five-, and six-year graduation rates for first-time-in-college (FTIC) [7], [8] and transfer students [12], [13]. Instructional complexity has not seen any significant adoption.

Our Project, Merging MIDFIELD with Curricular Analytics

We are currently in the first year of a project addressing the framework's potential to bring a renewed perspective to the MIDFIELD dataset. We are exploring four interlocking research questions in this project. The first two questions are descriptive, outlining how curricular complexity can vary across strata of interest to engineering education research and the BPE program. We will be capturing the curriculum as codified and as experienced by the student. Each of our research questions are posed generally, but our project seeks to focus on the experiences of underrepresented students.

RQ1: How does the complexity of the codified curriculum vary among institutions, disciplines, and matriculation models?

RQ2: What are the course-taking trajectories for different student populations (e.g., FTIC, changing majors, transfer)? And how do these student-enacted curricular maps compare through the lens of curricular complexity?

Capturing the curriculum as experienced by the student provides insight into how students traverse the curriculum differently. Moreover, transfer pathways can become more complicated than starters under certain conditions, like losing credits for core courses in the curriculum. We contend using the curricular complexity metrics provides a quantitative measure to compare these student experiences.

We are exploring five disciplines: Mechanical, Electrical, Chemical, Civil, and Industrial Engineering. These five majors were chosen due to the size of their enrollment and their prevalence among U.S. engineering programs.

A notable product of the analysis for research questions 1 and 2 will be a new dataset tied to MIDFIELD, a set of curriculum maps that can be readily imported into R and explored using the midfieldr package (or any other statistical computing platform). These maps will be broadly available to the research community upon completion of the project.

Next, we pose a comparative question. RQ3: To what extent do students follow the curriculum as codified? How does that vary by institutions, disciplines, matriculation models, populations, and pathways?

We will leverage these curricular maps to explore the extent to which students follow the curriculum as codified in university catalogs. For example, which populations are more or less likely to take courses as prescribed versus adapting plans of study to suit personal interests or needs? Who is retaking courses and/or transferring courses in from other institutions? These questions allow us to interrogate previous findings in the literature using curricular complexity as a curricular accessibility framework. Finally, our fourth question is correlational and concerns student outcomes.

RQ4: How is the curricular complexity experienced by students related to overall GPA, discipline stickiness, migration yield, and time-to-graduation?

We will correlate curricular complexity with established metrics within the community like overall GPA, persistence, and time-to-degree in addition to ecosystem metrics through MIDFIELD-associated work like stickiness and migration yield [14]. Stickiness refers to the percentage of students who ever enroll in a specific discipline and graduate in the same discipline. Migration yield is the normalized gain of migrating students that a discipline attracts and graduates within six years to the number of migrating students it could attract.

Thus, our project goals are:

1. To produce a new dataset linked to MIDFIELD, a set of curriculum maps, which will be publicly available to the research community.

2. To explore and understand the curricular complexity of engineering programs in five disciplines (Mechanical, Electrical, Chemical, Civil, and Industrial Engineering) across institutions and matriculation models.
3. To examine and compare the experiences of underrepresented students in navigating the curriculum and quantify their student-enacted curricular maps using curricular complexity metrics.
4. To investigate the extent to which students follow the codified curriculum, and how these behaviors vary by institution, discipline, matriculation model, population, and pathway.
5. To examine the relationship between curricular complexity experienced by students and their academic outcomes, such as overall GPA, discipline stickiness, migration yield, and time-to-graduation.

Year One Activities

Data Collection. Our first year was primarily focused on collecting data to address RQ1 and RQ3. We focused our search for curricular information to five disciplines: Civil Engineering, Electrical Engineering, Mechanical Engineering, Chemical Engineering, and Industrial Engineering at 13 MIDFIELD institutions with data current up to students entering in 2015 with six years of available data. For each institution, we collected the previous ten years of curricular data prior to the institution's last record in the appropriate format for network analysis. Given the parameters of our search, the upper bound for our data collection was 650 plans of study. However, each institution does not necessarily offer all five disciplines of interest nor for all years of consideration. Accounting for these practical gaps, our dataset contains 494 networks.

Planned Analysis. Now that we have the plans of study entered in a standardized format, we are currently conducting descriptive analyses across our strata of interest. Our first set of comparisons is between institutions and disciplines by plotting the trends in complexity over time and using box plots to characterize the variation across strata. We are taking this opportunity to verify data are entered correctly by examining cases where complexity did not change at all over the span of at least three years and suddenly jumped or declined by a considerable margin.

To address RQ2 and RQ4, we will be using association analysis to construct course-taking trajectories from the course table in MIDFIELD. A proof of concept for this idea is offered by Wang [15], who used data mining techniques to understand course-taking patterns of successful transfer students in the Beginning Postsecondary Students Longitudinal Study dataset. Association analysis allows us to find common combinations of courses that make up the broader trajectories that students follow when pursuing their degree. To build the individual trajectories, we will use Slim's algorithm [8]. The algorithm allows us to generate the associated network of courses by arranging them in the appropriate term. Slim's algorithm does not assign the exact prerequisites and corequisites; therefore, we will filter back through with our existing plan of study data to generate the necessary prerequisites and corequisites.

Once the trajectories are complete, we will calculate the structural complexity metrics that each student experienced by incorporating retaking behavior and major switching. These values will be clustered and disaggregated across strata, such as FTIC versus transfer, race, gender, and first-generation status. We plan to correlate structural complexity with ecosystem metrics like discipline stickiness and migration yield. By disseminating these results directly to institutional stakeholders and the broader engineering education community, we anticipate that this project can inform curricular design for all engineering students and help us understand what academic policies are inhibiting degree attainment for diverse groups.

Year One Achievements

Goal 1 of the project, which was to produce a new dataset linked to MIDFIELD that contains curriculum maps for each major, has been achieved. The project team has successfully compiled 494 plans of study. The data analysis phase is currently in progress, and the results will be used to address our four research questions.

Related to Goal 1, the project team cataloged common questions asked by undergraduate research assistants during data collection as a measure of quality and transparency for research processes. The outcome will be an FAQ that can be used for broader use, including replication of the project, in the future. By documenting and answering frequent questions, the project team hopes to provide a comprehensive resource for others interested in replicating the project or using the data and results produced by the project.

To scale the analyses, we have written an alpha version of an R package which is used to calculate the curricular complexity metrics. It has undergone validation by reproducing results from a previous effort. Completing the R package is a significant milestone in the project because it provides a platform for researchers to easily import and explore the curriculum maps we produced. Moreover, the validation of the package provides additional confidence in the accuracy and reliability of the results later in the project.

Changes to Our Approach

In the original conceptualization of the project, it was planned that all MIDFIELD institutions would be incorporated into the plan of study data set in anticipation of being compared to actual course-taking trajectories by students. Upon reviewing the available data within the most recent release in collaboration with the data steward, we identified seven institutions (i.e., North Carolina Agricultural and Technical University, Florida A&M University, University of Florida, Georgia Institute of Technology, University of Utah, and Virginia Tech) that had not submitted new data to MIDFIELD since 2009.

Moreover, reliable catalog information needs to be retrieved to make appropriate comparisons between student course-taking trajectories and the codified plans of study. As we go back to retrieve catalog information on older websites, accessible prerequisite/corequisite and plan of study information becomes sparser. There is considerable variance across institutions regarding their approach to storing catalogs, even for catalogs within the last decade. Thus, it becomes more difficult to compile the necessary parameters for the network models accurately.

Given these two considerations, it became more financially prudent to scope the data collection to institutions with more recent data updates. When the Undergraduate Research Assistants were tasked with fetching curricular-related data earlier than 2010, significant effort was being expended to locate catalogs using the institution's native search functions and the Wayback Machine. To avoid project delays, we retained the longitudinal element of our design by starting at the most recent year for each institution and looking back a full decade. The seven institutions with data all before 2009 were excluded for the time being, unless it becomes more salient what value add could be achieved by applying person-hours to the endeavor. With our refocused scope, the data collection was completed on schedule.

Future Work and Potential Roadblocks

Under the supervision of the PhD student and the PI, the undergraduate students will be responsible for performing descriptive analyses of the data collected during the next reporting period. The work will involve calculating the various metrics necessary for comparing disciplines and institutions. These results advance us toward addressing research questions RQ1 and part of RQ2.

At the same time, the PhD student will be conducting association analysis to form course-taking trajectories over the relevant decade of data collected. These course-taking trajectories will be drawn from the MIDFIELD dataset to compare them to the codified plan of study data collected during this reporting year. The focus of the association analysis will be on answering the second, third, and fourth research questions, which are centered around the extent to which students follow the codified curriculum and the relationship between curricular complexity and student outcomes. These trajectories will be disaggregated among the strata of interest: matriculation models, institutions, disciplines, populations, and pathways.

After discussing our next steps forward, we identified two cases where some additional care may be needed to make appropriate conclusions and comparisons. First, we anticipate that the analysis of course-taking data may be challenging in cases where students have enrolled in more than one major. Students who pursue multiple majors often have unique pathways through the curriculum that are unlikely to map cleanly to a specific codified plan of study that we have collected. We had already considered students who switch majors as a stratum of interest, but it may be necessary to analyze the subset of students who double major separately as their own strata as well.

Second, students transfer credits from other institutions (both double majoring students and FTIC students), which is not readily obvious in the MIDFIELD dataset. We also plan to use the attempted credit hours and earned credits hours as proxies for transfer pathways to recognize students who transferred substantial amounts of credit from other institutions despite not being traditional transfer students.

Broader Impacts

Training for Undergraduate and Graduate Students. This project has provided training under the supervision of the PI for a PhD student in Engineering Education, who plans to complete her dissertation on one dimension of this project. She is being trained on the statistical programming platform, R, to conduct association and network analyses. Moreover, a team of five undergraduate research assistants has been provided with opportunities to engage in engineering education research. Four undergraduate engineering students were hired to assist with data collection in Fall 2022, in addition to another undergraduate student who was already assisting the PI with a related unfunded project about curricular complexity. These students were trained in the responsible conduct of research, exposed to the ambiguities and challenging decisions analysts must make during data collection and revising the sampling strategy within the research design, and collaborated digitally using the features in Microsoft Teams.

One undergraduate student began attending the PI's weekly graduate research group meeting and has expressed interest in contributing more to the project and others beyond this effort. The retention of these Undergraduate Research Assistants was unanticipated but welcomed, with four of the five continuing into Spring 2023 to further engage with the research. In Spring 2023, the Undergraduate Research Assistants engaged with the basics of qualitative research as we reviewed our processes from the previous semester and generated a method for others to replicate our research design. They also learned to apply descriptive statistics to the data they collected to spot potential errors and outliers and later the process of generating new quantitative metrics to address preexisting and unexpected research demands.

We believe there is a considerable opportunity to continue engaging three undergraduate students in research as the project continues. The fourth graduated in Summer 2023.

Applying Curricular Analytics Longitudinally. To our knowledge, there has been no systematic longitudinal study of curricular complexity in engineering or other disciplines using the Curricular Analytics framework by Heileman and colleagues. Accordingly, several ambiguities emerged during data collection that required deliberations within the research team to decide the prudent steps forward. The original authors do not provide guidance regarding how to process data beyond the most rudimentary situations. For example, the data entry for conventional plans of study assumes that prerequisite relationships do not contain complex conjunctions (i.e., MATH 101 and MATH 102, OR MATH 101H), which occurs frequently enough to be an issue. To keep track of questions during our data collection processes, we maintained a Microsoft Teams channel that included the PI, PhD student, and all undergraduate research assistants. After data collection was complete, all chats from the channel were exported and coded to find common questions. These questions were then categorized into a frequently asked questions document that is being incorporated into an R package that will be made available to the broader community. A separate publication is available detailing these questions.

A New Standalone and Companion R Package for MIDFIELD. The R package developed in this project will be paired with the existing "midfieldr" [16] and "midfielddata" [5] packages. The "midfieldr" package is a set of functions that enables researchers to streamline the analysis of student record data within MIDFIELD, whereas the "midfielddata" package provides toy data for those interested in conducting analyses using the full dataset can understand its format and experiment with "midfieldr" functionality. Together, the R package and the "midfieldr" package will enable researchers to:

- Access and analyze a large dataset of course-taking information from multiple institutions and disciplines.
- Explore the curricular complexity of different programs and disciplines.
- Visualize the course-taking experiences of students over time and across different pathways.
- Correlate curricular complexity with student outcomes such as overall GPA, persistence, and time-to-degree.

By pairing these two packages, researchers will be able to gain a deeper understanding of curricular complexity and the experiences of students in higher education. This will help to inform the development of better educational policies and practices that support student success and improve the quality of engineering education. Alpha versions of the package will be made available upon request, with the intention of a broader release by the project's close. Requests for alpha access can be made to the PI via email.

Informing Curricular Design for All Students. Understanding curricular complexity can help to streamline programs for students and keep them on track for graduation. By analyzing the curriculum as experienced by students, researchers can gain insight into the challenges and obstacles that students face as they progress through their program – such as identifying non-obvious bottlenecks. This information can be used to identify areas where the curriculum is unnecessarily complex or where students are at risk of deviating from the prescribed path. For example, by examining the experiences of different groups of students, institutions can identify areas where additional support and resources are needed to help them succeed. This may involve making changes to curricular policies, providing additional academic advising and support, or developing new programs and initiatives to support student success.

Conclusion

We anticipate this work being useful to researchers and practitioners interested in systematic analyses of curricula across disciplines, especially in combination with student data to explore retention-related issues for FTIC students. The dataset we created will be freely available, and others are encouraged to add their own plans of study. We offer the FAQ [17], along with data conventions and the R package, as resources to best facilitate the large-scale analysis of this type of network data. As the dataset grows, we anticipate the ability of the community to understand and interrogate the programmatic barriers to student success in engineering across the nation will also expand – leading to a cornucopia of previously unexplored questions at scale.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. BPE- 2152441. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] F. Curry and J. DeBoer, "A Systematized Literature Review of the Factors that Predict the Retention of Racially Minoritized Students in STEM Graduate Degree Programs," in *2020 ASEE Virtual Annual Conference Content Access Proceedings*, Virtual On line: ASEE Conferences, Jun. 2020, p. 34069. doi: 10.18260/1-2--34069.

[2] E. A. Kuley, S. Maw, and T. Fonstad, "Engineering Student Retention and Attrition Literature Review," *Proc. Can. Eng. Educ. Assoc. CEEA*, Aug. 2015, doi: 10.24908/pceea.v0i0.5813.

[3] A. Sithole, E. T. Chiyaka, P. McCarthy, D. M. Mupinga, B. K. Bucklein, and J. Kibirige, "Student Attraction, Persistence and Retention in STEM Programs: Successes and Continuing Challenges," *High. Educ. Stud.*, vol. 7, no. 1, pp. 46–59, 2017.

[4] M. W. Ohland, R. A. Long, S. M. Lord, M. K. Orr, and C. E. Brawner, "Expanding Access to and Participation in the Multiple Institution Database for Investigating Engineering Longitudinal Development," presented at the 2016 ASEE Annual Conference & Exposition, Jun. 2016. Accessed: Feb. 12, 2023. [Online]. Available: https://peer.asee.org/expanding-access-to-and-participation-in-the-multiple-institution-database-for-investigating-engineering-longitudinal-development

[5] R. Layton, R. Long, M. Ohland, M. Orr, and S. Lord, "midfielddata: MIDFIELD data sample." 2022. [Online]. Available: https://midfieldr.github.io/midfielddata/

[6] G. L. Heileman, C. T. Abdallah, A. Slim, and M. Hickman, "Curricular Analytics: A Framework for Quantifying the Impact of Curricular Reforms and Pedagogical Innovations," *ArXiv181109676 Phys.*, Nov. 2018, Accessed: Aug. 04, 2021. [Online]. Available: http://arxiv.org/abs/1811.09676

[7] D. M. Grote, D. B. Knight, W. C. Lee, and B. A. Watford, "Navigating the Curricular Maze: Examining the Complexities of Articulated Pathways for Transfer Students in Engineering," *Community Coll. J. Res. Pract.*, pp. 1–30, Aug. 2020, doi: 10.1080/10668926.2020.1798303.

[8] A. Slim, "Curricular Analytics in Higher Education," Dissertation, The University of New Mexico, 2016. Accessed: Feb. 24, 2023. [Online]. Available: https://www.proquest.com/docview/1873863748?pq-origsite=gscholar&fromopenview=true

[9] J. Wigdahl, G. L. Heileman, A. Slim, and C. T. Abdallah, "Curricular Efficiency: What Role Does It Play in Student Success?," presented at the 2014 ASEE Annual Conference & Exposition, Jun. 2014, p. 24.344.1-24.344.12. Accessed: Jan. 20, 2023. [Online]. Available: https://peer.asee.org/curricular-efficiency-what-role-does-it-play-in-student-success

[10] D. Waller, "Organizational factors and engineering student persistence," Dissertation, Purdue University, 2022. [Online]. Available: https://doi.org/10.25394/PGS.21606342.v1

[11] G. L. Heileman, W. G. Thompson-Arjona, O. Abar, and H. W. Free, "Does Curricular Complexity Imply Program Quality?," presented at the American Society for Engineering Education, Tampa, Florida, 2019, pp. 1–13.

[12] D. Reeping and D. Grote, "Rethinking the Curricular Complexity Framework for Transfer Students," presented at the ASEE Virtual Annual Conference Content Access, Virtual, 2021, pp. 1–24. doi: https://peer.asee.org/37680.

[13] D. Reeping and D. Grote, "Characterizing the Curricular Complexity Faced by Transfer Students: 2+2, Vertical Transfers, and Curricular Change," presented at the American Society for Engineering Education Annual Conference, Minneapolis, MN, 2022, pp. 1–16. doi: https://peer.asee.org/41462.

[14] S. M. Lord, M. W. Ohland, R. A. Layton, and M. M. Camacho, "Beyond pipeline and pathways: Ecosystem metrics," *J. Eng. Educ.*, vol. 108, no. 1, pp. 32–56, 2019, doi: 10.1002/jee.20250.

[15] X. Wang, "Course-Taking Patterns of Community College Students Beginning in STEM: Using Data Mining Techniques to Reveal Viable STEM Transfer Pathways," *Res. High. Educ.*, vol. 57, no. 5, pp. 544–569, Aug. 2016, doi: 10.1007/s11162-015-9397-4.

[16] R. Layton, R. Long, M. Ohland, M. Orr, and S. Lord, "midfieldr: Tools and methods for working with MIDFIELD data in 'R.'" 2022. [Online]. Available: https://midfieldr.github.io/midfieldr/

[17] D. Reeping, S. Padhye, N. Rashedi, "A Process for Systematically Collecting Plan of Study Data for Curricular Analytics," presented at the American Society for Engineering Education Annual Conference, Baltimore MD, 2023, pp. 1–19.