EFFECTIVE LABORATORY EDUCATION WITH TEXTILE: TUTORIALS IN EXPERIMENTALIST INTERACTIVE LEARNING

Hawley Helmbrecht and Elizabeth Nance
University of Washington • Seattle, WA 98195

INTRODUCTION

The chemical engineering core undergraduate and graduate curricula provides excellent instruction in numerical methods and fundamental concepts of thermodynamics and transport underlying biological and physical systems. Additionally, the chemical engineering field is rapidly being transformed by high-throughput data and data science methods for organizing, processing, and analyzing large data sets.[1] Discussion regarding data science integration in chemical engineering is ongoing, and some chemical engineering programs including – but not limited to – the University of Washington, Purdue University, Columbia University, and Georgia Tech offer data science options, certificates, or concentrations.[2] Chemical engineering students at the undergraduate and graduate levels have a unique opportunity in a range of research areas to apply and utilize data science methods for handling complex systems and data.[3] This is especially true in chemical engineering applications to biological research since fields involved in generating or analyzing biological data are developing methods and software for handling high-modality and high-value data sets.[4] Chemical engineering research laboratories are making headway in biologically-relevant technologies such as biosensors, biomaterials and tissue engineering, protein engineering, and nanomedicine. Therefore, chemical engineering students engaging in biologically-focused chemical engineering research labs have opportunities to use or develop data science techniques, creating a powerful training environment for the next generation of chemical engineers to handle high-throughput and high-modality biological data.

Student experiences in STEM education are greatly affected by participation in research.[5, 6] Undergraduate student involvement in chemical engineering research introduces students to cutting-edge research and increases knowledge and skills in critical thinking, problem solving, conducting research projects, and working independently.[5] Additionally, research experience utilizing data science introduces students to skills in managing large data sets, scaling computational processes, and utilizing statistics for data analysis.[7] Finally, student research experiences in lab have been promoted as a method for improving student outcomes,[8, 9] increasing retention,[10] and developing student-mentor relationships.[11-13]

However, the ability of most labs to teach and train undergraduate and high school students is greatly limited. Some factors limiting the number of student trainees a research lab can manage include the number of graduate students, post-doctoral scholars, and research assistants in the lab,[14] personnel time and capacity for mentorship and training among other responsibilities,[15] faculty availability and willingness for mentorship, university reward structures for mentoring,[16] and faculty interest in increasing diversity in academia.[17] Additionally, research lab turnover for expertise and training is relatively fast, at 2-6 years for gradu-

__Hawley Helmbrecht__ is a PhD candidate in the Department of Chemical Engineering at the University of Washington (UW) completing the advanced data science option. Her research interests focus on developing image processing pipelines for quantifying brain cell morphology. She has served as a teaching assistant for the senior capstone design and mass and energy balance courses in UW Chemical Engineering and mentored fourteen students in her research area. ORCID: 0000-0002-4797-0130

__Elizabeth Nance__ is the Jagjeet and Janice Bindra Career Development Endowed Associate Professor in Chemical Engineering at the University of Washington (UW), Associate Professor of Bioengineering, and adjunct in Radiology and the eScience Institute. She serves as the Associate Chair for Undergraduate Studies in Chemical Engineering at UW and founded Women in Chemical Engineering (WChE) in 2016. Her research focuses on technology development for the treatment of pediatric and brain-related diseases. ORCID: 0000-0001-7167-7068

ate students to learn techniques and graduate, and ranging from <1 to up to 4 years for undergraduate students. Graduate students may master a technique in a few years, train other personnel for a year, and then subsequently graduate, leaving a new student to be the teacher. Conversely, lab training may occur at irregular intervals, on an individual trainee basis, and with differing levels of starting experience. The variable nature of independently training new researchers in laboratories increases the time burden of training for every individual technique and for each unique tailoring of the training approach for specific researchers' current expertise.

To increase the number of students who can participate in high school and undergraduate research opportunities in our research lab, we developed Tutorials for EXperimentalisT Interactive LEarning (TEXTILE). TEXTILE is a formal curriculum for laboratory training that is designed to decrease the time burden for training new students, thus increasing the number of students a research lab can mentor. We showcase our lab's application of TEXTILE in chemical engineering-based neuroscience research. TEXTILE connects the structured concept familiar to most students of a defined curriculum with support staff to the open-ended nature of research projects in laboratories.[18]

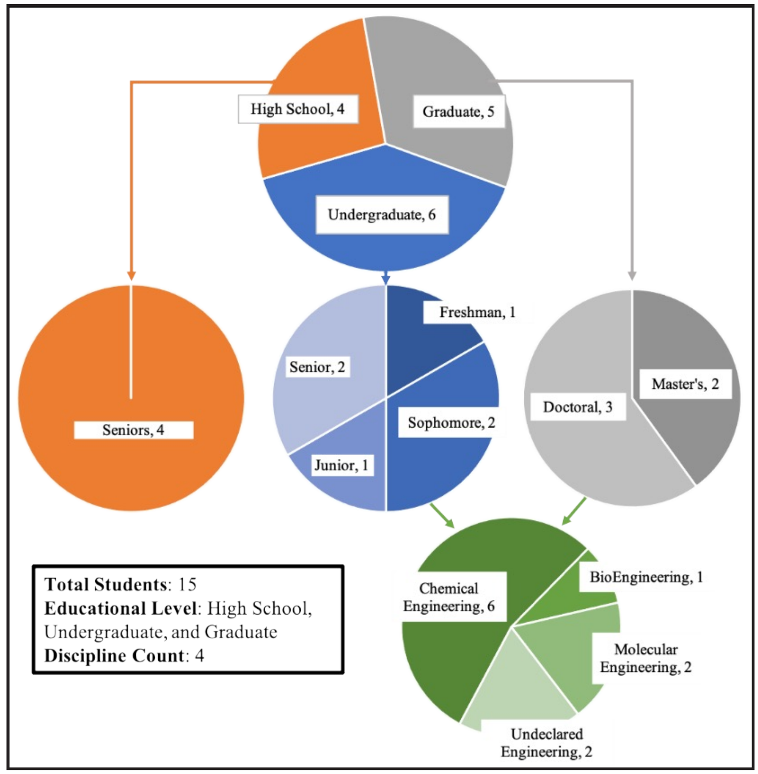

Figure 1. *Student Population: Student breakdown by overall education level, year in education level, and degree program if applicable.*

METHODS

Study Population

The study population for the TEXTILE curriculum includes fifteen students: four high school students, six undergraduates, and five graduate students (Figure 1). The relatively even split of educational levels shows the versatility of TEXTILE as an expertise-independent platform. Of the students in the TEXTILE program who were fully enrolled in a specific degree program, over 50% were chemical engineering students and the rest were in other engineering majors or undeclared engineering freshman. Of the students participating, half the graduate students, three of the high school students, and half the undergraduate students had little to no coding experience. The other graduate students with coding experience had completed an introductory course in data science and software engineering offered through the data science option at the University of Washington. The undergraduates with coding experience varied in their level of expertise, with two students who had taken an introduction to Python™ course and one student with an intermediate skill level in Python. Finally, only one high school student had coding experience in languages other than Python at the beginning of the TEXTILE curriculum. All students were participating in summer research with the Nance Lab in the Department of Chemical Engineering at the University of Washington. All activities necessitated online interaction due to restrictions in place during the COVID-19 pandemic.

Developing a Lab-Based Curriculum

The TEXTILE curriculum teaches data science methodologies within the context and practices of a specific laboratory. As our educational goal, we chose a common methodology in the lab with high applicability to various experimental methods – brain cell segmentation with image processing – such as our analysis completed on microglia, the brain's immune cells, in a rat brain slice model.[19] With an educational goal set, the data science components involved in reaching the cell segmentation goal were broken down into digestible pieces; the image processing curriculum series includes Images as Numbers, The Power of Thresholding, and Segmentation/Labeling Cells/Getting Quantitative Cell Characteristics. Once the methodology specific pieces were determined as module goals, we identified basic data science methods needed to support the image processing educational series. The data science series was broken down into four components: A Brief History of Coding, A Data Scientist's Role, GitHub and Version Control, and Intro to Data Management. Finally, an Intro to Python module was developed for students with no coding experience that provides resources and activities in exploring methods for learning Python.

Any presentations were developed using Microsoft Office PowerPoint® or Canva® and made readily accessible through Google Drive®. The presentations were also uploaded to the TEXTILE GitHub® repository hosted on our lab GitHub account.[20] All data science and image processing modules were uploaded to the TEXTILE GitHub as Jupyter Notebooks®. Pre-activity modules were hosted as their own notebooks in the repository while post-activity modules were included at the bottom of all main module Jupyter Notebooks.

Data Collection

Data collection occurred in multiple formats through feedback forms, verbal feedback, GitHub repositories, and through in-session Zoom® video conferencing and Slack® Technologies, Inc. emoji-based reactions. We developed feedback forms for the first few modules using Google Forms®. The feedback forms included questions with scalable answers about the adequacy of time available for each portion of the module, clarity of instruction, quality of activities, and student interest in the major portions of the module. The feedback form also included short answer questions about the favorite and least favorite part of each module. Verbal feedback was obtained during the main module during check-ins throughout the presentation. Most verbal feedback pertained to pacing, accessibility of code, and visibility of instructor materials. GitHub repositories were created by students in the GitHub and Version Control module on the data science specific learning path. These repositories and student comments were pushed to our lab GitHub for review by instructors. Finally, Zoom video conferencing and Slack Technologies, Inc. emoji-based reactions were utilized to gather quick student response and interaction during modules. An example of Zoom video conferencing emoji use for student interaction is selecting a green checkmark versus a red 'x' during the Data Scientist's Role module to convey student perception of data scientists. Emoji-based reactions in Slack Technologies, Inc. supported pre-module activity and post-module activity check-ins for quality of the activity and student participation – if students had questions or enjoyed the activity, they could respond to a Slack message with specific emojis.

EDUCATIONAL METHODOLOGY

We developed TEXTILE as a learning platform for wet-lab and data science-integrated research. We specifically applied the platform to data-science enabled segmentation of fluorescent images of brain cells. Since TEXTILE is a module-based, semi-linear educational system developed with laboratory-based learning in mind, modules combine data science fundamentals with specific biological imaging processing techniques. For example, we pair a data science module about data management with an image processing module on quantitative cell morphology characterization. The modular format allows us to train students on managing experimental data and pair that with quantifying features of cells that are relevant to their specific project of interest. Additionally, we designed TEXTILE such that the learning process is adaptable to a beginner's skill level. Learning paths are based on entry knowledge but allow students to pick and choose their modules based on interest or previous experience. For example, someone with ample coding knowledge can skip the introductory Python modules and begin with the experimental image processing modules. Finally, we designed TEXTILE for a variety of degree levels. As a laboratory at a large public university, we engage with undergraduate and high school students, in addition to master's thesis and doctoral dissertation students. We accomplished degree-level accessibility through the semi-linear system that assumes no prior knowledge before beginning the TEXTILE program.

Textile Outline

TEXTILE is composed of topic-specific modules. Each module is designed around a data science topic, lasts an hour and a half, and contains four sections: the pre-module activity, main module, post-module reflection, and feedback form. The instructor sends the pre-module activity to students for self-driven learning a week before teaching the main module. The pre-module's purpose is to prime students for the module's main lesson with a high-level activity. For example, in our data management module, students analyze a grocery shopping trip and decide where to store specific categories of food and perform a space evaluation on how much room in each storage location the food will take up.

The main module aims to teach the bulk of material in digestible lectures broken up by interactive sessions. The main module's detailed structure consists of six parts: introduction presentation, interactive session one, main presentation workshop, interactive session two, concluding presentation, and a reflection session (Figure 2). The introduction presentation introduces the module's topic, outlines the structure the main module will follow, and leads into the first interactive session. Interactive session one guides the student through an exploration of the module material. For coding-based modules, instructors lead interactive sessions with Jupyter Notebook. During interactive session one, the instructor walks students through the code's basics as students code along. After the students have written the code, the instructor begins the main presentation, which has two purposes: (1) to increase the students' connection between the pre-module activity and the module introduction and (2) to dive deeper into the technical aspects of the first interactive

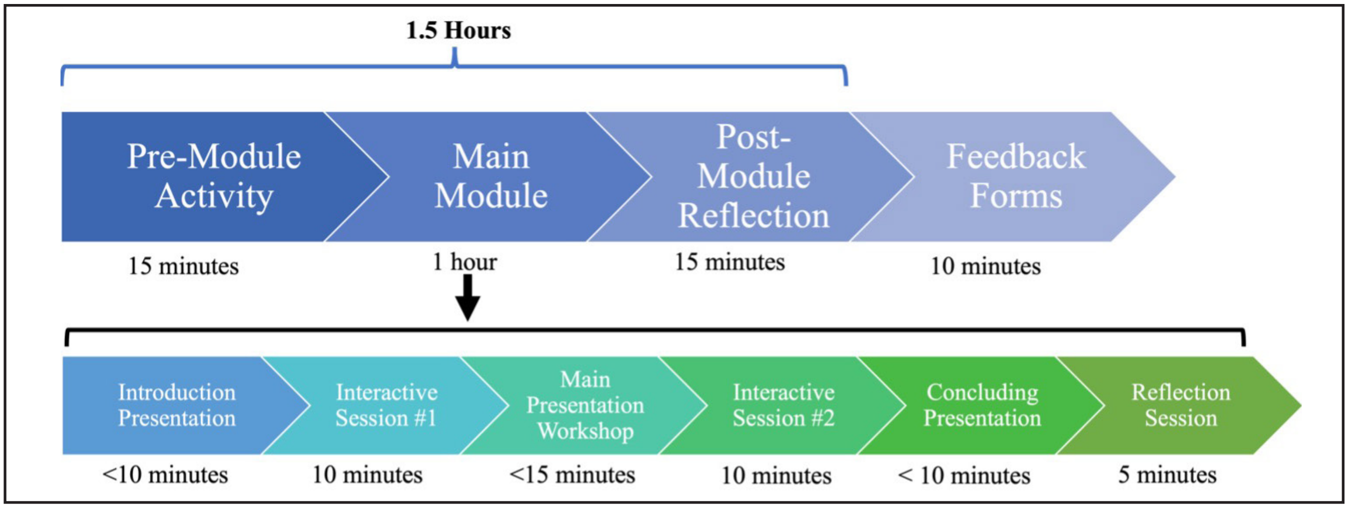

Figure 2. *TEXTILE Methodology. The top row of the figure indicates the time length of the four parts of the TEXTILE process. The bottom row of the figure indications the six sections of the Main Module and the time length of each section.*

session. Once students understand the technical elements of the main module, they begin interactive session two, where they apply technical details of the lesson to further explore the code written in the earlier interactive session. After students have an opportunity to explore, the instructor gives the concluding presentation, which includes wrapping up the main takeaways, emphasizing the module's motivation, and thanking students for their participation.

Course Design and Curriculum

When developing TEXTILE, we intentionally designed both the module system and the module curriculum to encompass specific purposes and utilize a variety of learning methods (Table 1).

The TEXTILE curriculum is a semi-linear process with three main educational pathways: two computational pathways – data science and image processing – and the research lab specific pathway (Figure 3). In the research lab specific pathway, we developed the modules for three purposes: tethering the other lessons to specific experimental techniques and goals of our lab, training other educators how to develop modules for TEXTILE, and applying the lessons learned in the other pathways to published data from our lab. Meanwhile, the data science pathway introduces students to coding, data science, and data management tasks related to experimental research. Finally, the image processing specific pathway deep dives into the methodologies utilized and developed in our lab for quantifying brain cell images.

Alongside the three pathways, there are four entry points for lessons: no coding experience entry, contextual lab specific entry, module development entry, and previous coding experience entry. With the no coding experience entry students begin with an Intro to Python module that tasks students with learning Python through various provided references that include video tutorials, online books, and interactive code learning website links. After the Intro to Python module, the no coding experience students then move through the entire data science and image processing pathways. With the previous coding experience pathway, students choose their beginning data science and image processing module based on their coding experience – some students choose to start at the beginning regardless of coding experience as

TABLE 1
Learning Methods and Purpose of Each Module Portion

Module Portion	Purpose	Methods
Pre-Module Activity	Prime the students for the lesson	Student led, inquiry-based learning, designing research questions, designing data management plans, literature reading
Main Module	Provide the main educational material	Teacher led, reflective discussion of pre-module activity, code-along, individual code exploration
Post-Module Activity	Refresh and reinforce the students' learning while diving deeper into individual interests	Student led, inquiry-based learning, reflective discussion

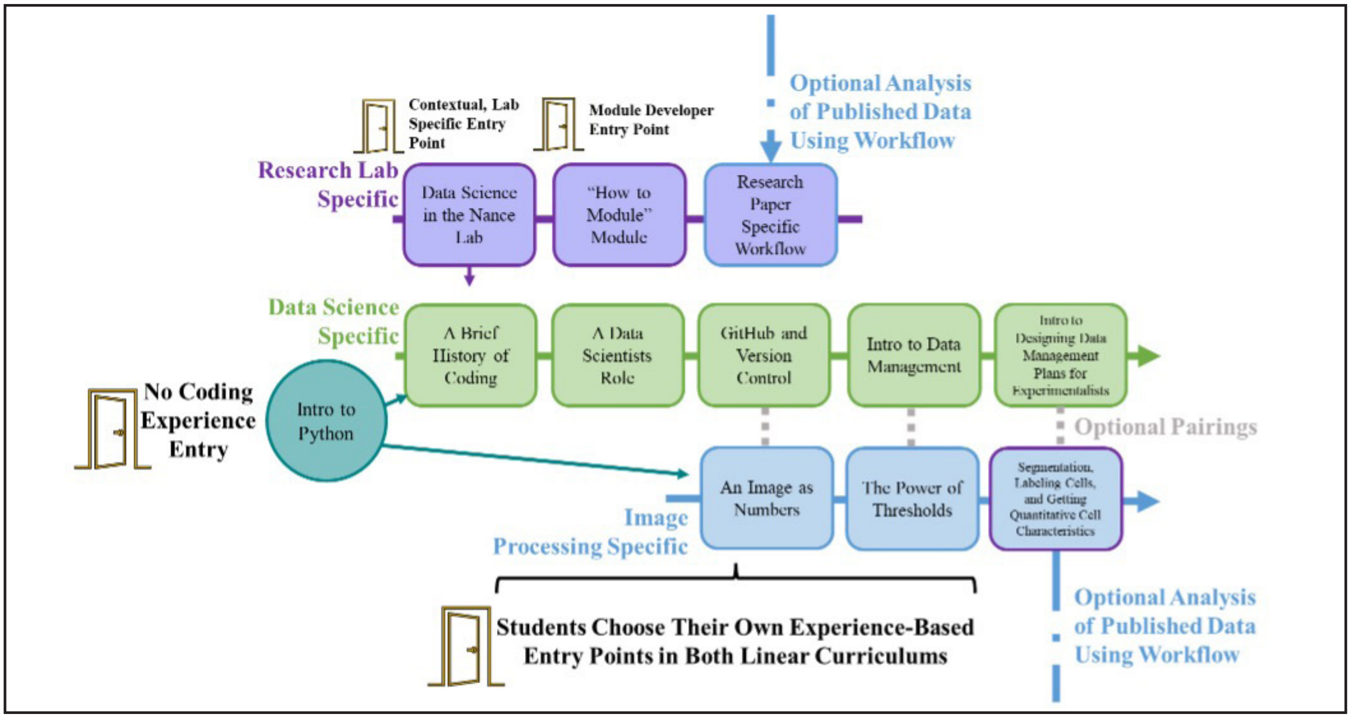

Figure 3. *TEXTILE Curriculum. Demonstrates the three educational paths - Research Lab Specific (top row), Data Science Specific (middle row), and Image Processing Specific (bottom row). Entry points are displayed as doorways.*

a refresher of their knowledge. All students that go through TEXTILE to begin working in our lab start with the contextual lab specific entry that provides background information about our lab's experimental work, data science research, and data science goals before beginning the data science or image processing pathways. Finally, those in instructional roles and interested in developing modules may begin with the "How to Module" module to expand the TEXTILE program.

Course Syllabus

There were eleven developed modules for TEXTILE during the course as described in this paper (Table 2). Each module is associated with one of the three specific pathways and serves a unique purpose in training students on the final goal of applying the Segmentation/Labeling Cells/Getting Quantitative Cell Characteristics module methodology to published data from our research lab.

Learning Objectives

Each module has at least three learning objectives that are split into the pre-module activity, main module, and post-module activity (Table 3). Learning objectives for each module are achieved in different ways based on the purpose of the module and concept. For coding intensive modules, the pre-module typically includes a conceptual learning objective that introduces the students to ideas taught in the main modules, while the main module focuses on coding-based outcomes as learning objectives. For data science modules without coding intensive sections, learning objectives typically revolve around conceptualizing data science methodologies within an experimental research example, such as those described in Intro to Designing a Data Management Plan for Experimentalists.

Course Structure

The course was offered in both synchronous and asynchronous formats. In the synchronous format, data science and image processing modules were offered once a week over a period of six weeks following the educational pathway that assumes zero coding knowledge. A week prior to the main module, students were sent the pre-module activity and asked to complete the activity and prep question or answers for the main module. The main module was hosted live over Zoom video conferencing by a graduate student conducting research related to the module in the lab. The post-module activity was given to students at the end of the main module while the feedback form was sent via email to students at least a day after the main module session to give them time to complete the post-module activity. In the asynchronous format, modules were pre-recorded or additional documentation was provided in Jupyter Notebooks to replace the role of a lecturer during modules. Students com-

TABLE 2 TEXTILE List of Modules and Descriptions		
Module Name	**Module Topic**	**Module Description**
Data Science Role in the Nance Lab	Data Science	Introduces the overall research done in the Nance Lab and how data science enhances insight gained in each of the wet lab areas.
Brief History of Coding	Data Science	Introduces students to the history of coding languages and the large variety and purposes of many coding languages.
A Data Scientist's Role	Data Science	A highly interactive lesson that encourages students to question what they consider data and begin to view themselves as data scientists.
GitHub and Version Control	Data Science	Introduces students to GitHub and version control for use in future lessons.
Intro to Data Management	Data Science	Introduces students to data management basics in both experimental and data science-based work.
Intro to Designing a Data Management Plan for Experimentalists	Data Science	Encourages students to develop a data management plan for a wet lab and data science combined experiment.
Intro to Python	Computer Science	Provides many resources for introductory and topic specific Python education.
Intro to Image Processing: An Image as Numbers	Image Processing	Encourages students to rethink what they know about images and introduces them to the numbers behind the images they are seeing with use of NumPy arrays.
The Power of Thresholds	Image Processing	Shows students the great power and possible bias in thresholding cell images.
Segmentation, Labeling Cells, and Getting Quantitative Cell Characteristics	Image Processing	Utilizes the lessons from Power of Thresholds to teach students how to segment cells and produce a .csv file of quantified cell morphology features from the cells segmented in previous lessons.
The "How to Module" Module	Informational	Teaches educators in the TEXTILE program how to outline and develop a module in a standard way while allowing for module specific flexibility and creativity.

pleted group discussion over Slack Technologies, Inc. and checked in regularly with a graduate student mentor about progress.

While the course was hosted virtually due to COVID public health concerns, in-person or hybrid options are worth pursuing in the future. Some benefits of a fully virtual curriculum are accessibility for trainees not within immediate commuting distance and easier automatic recording of lectures. For the use of TEXTILE in this paper, the virtual option worked well since all curriculum was data science based. However, benefits of a hybrid or in-person option include the ability to do laboratory tours or demos of wet-lab procedures for data acquisition. Additionally, hybrid/in-person options can allow students to experience work in a full-time research lab more so than focusing on online time specifically for the course.

RESULTS: STUDENT RETENTION IN LABORATORY RESEARCH

TEXTILE is a researcher training program developed for the Nance Lab that emphasizes the benefits of data science, computer science, and chemical engineering in medical research. Although career interests were not specifically explored through the program, many students expressed significant impact of the curriculum on both their undergraduate program and future career goals (Table 4).

Our overall educational purposes of TEXTILE presented in this manuscript – beyond maintaining student research positions during COVID-19 induced laboratory shutdown – were to encourage students to continue doing research in our lab after participation and to increase student comfort with data science. While there is no control group to directly

TABLE 3
Individual Sub-Module Learning Objectives

Module Name	Pre-Module	Module	Post-Module
Data Science Role in the Nance Lab	• Reflect on student goals with TEXTILE • Familiarize students with research lab website	• Connect the major research lab focus areas together • Verbalize the main use of data science in the lab	N/A
Brief History of Coding	• Identify two major coding languages • Explore language similarities and differences	• Recognize the wide diversity of coding languages • Place languages on timelines	• Write print statements and visualize circles in multiple coding languages
A Data Scientist's Role	• Explore current understanding of a data scientist	• Encounter the diversity of data and flexibility of the data scientist role	• Reflect on how the idea of a data scientist may have changed
GitHub and Version Control	• Open GitHub student accounts	• Clone the TEXTILE repository • Push code-along results	• Identify important aspects of GitHub repositories
Intro to Data Management	• Identify variables, data sharing options, and basic experimental methodology	• Apply data management concepts to imaging data	• Apply data management principles to research data from lab
Intro to Designing a Data Management Plan for Experimentalists	• Explain definition and purpose of a data management plan	• Apply data management concepts to imaging experiments	• Apply data management principles to research projects in lab
Intro to Python	• Explore Python learning options and methods	• Write: print, import, and visualization statements in Python and Jupyter Notebook	• Continue exploring Python with additional learning options and methods
Intro to Image Processing: An Image as Numbers	• Import .tiff images to Python • Visualize .tiff images in Jupyter Notebook	• View images as numerical arrays • Split color channels on fluorescent cell images	• Load an image of choice into Python and view as numerical array
The Power of Thresholds	• Learn the definition of a threshold • Apply the concept of thresholds in non-scientific and scientific applications	• Apply thresholds with Python on fluorescent cell images	• Explore cell thresholding in literature with provided resources
Segmentation, Labeling Cells, and Getting Quantitative Cell Characteristics	• Troubleshoot and reflect on previous or reoccurring errors	• Run the whole cell quantification pipeline on a single fluorescent cell image • Produce a .csv of quantified cell features	• Visualize morphological features of cells from the .csv using Python, Excel, or other graphing software
The "How to Module" Module	N/A	• Learn the formatting and structure of TEXTILE	• Develop a new module

TABLE 4 Student Retention and Career Interests after TEXTILE					
Student Descriptor	**Students Retained in the Lab**	**Students Retained in Engineering**	**Students Retained in Chemical Engineering**	**Students Continuing in Data Science Projects**	**Career Interests after TEXTILE**
High School	75% (3)	75% (3)	N/A	50% (2)	• Graduate Degree (Medical) (1) • Computer Scientist (1) • Chemical Engineer in Industry (1)
Undergrad	100% (6)	100% (6)	67% (4)	67% (4)	• Chemical Engineer in Industry (2) • PhD Chemical Engineering (2) • PhD Computer Science (1) • Medical School (1)
Graduate	100% (4)	100% (4)	75% (3)	50% (2)	• Academia (2) • Data Scientist Chemical Engineering Field (2)

compare TEXTILE students due to complications such as a national pandemic and no comparison cohort of the same size/education levels entering the Nance Lab previously, we still find value in student retention outcomes regarding the diversity of projects that stemmed from this curriculum, the transition of students to both wet-lab and data science projects, and the career considerations of our students. We outline the wide variety of projects and career trajectories that were impacted by the TEXTILE curriculum to show the flexible applications of TEXTILE to meet specific lab needs and educational goals.

High School Student Trainees

Of the four high school students who underwent the TEXTILE curriculum and decided to continue in the lab, two are undergraduates at the University of Washington in the pre-engineering track and a third is enrolled in a local two-year college. For the two students continuing data science research projects in the lab, we worked with each students' individual interests to determine an optimal project. The student interested in a graduate degree in medicine will be applying the image processing techniques from TEXTILE to neonatal disease models developed in our lab in collaboration with the Neonatal Neuroscience Lab at the University of Washington, while the student interested in a computer science career is developing automated data visualization of the cell morphology features resulting from the image processing pipeline taught with TEXTILE. The third student – interested in being a chemical engineer in industry and exploring wet-lab research – is training in confocal microscopy to develop a high-throughput methodology for localized magnification of cells for analysis with the image processing pipeline taught in TEXTILE.

Undergraduate Student Trainees

Of the undergraduate students, all students continued research in the lab and in engineering majors. However, two of the students chose or were already enrolled in non-chemical engineering majors – one bioengineering and the other computer science. Additionally, two students chose to continue in the lab with non-data science projects. Of these two students, one student has since graduated from UW Chemical Engineering and is now working in industry. The student expressed that the program gave them a chance to learn Python to enhance their skillset and resume. The second undergraduate officially joined the lab as a research staff member, and although they are not currently applying the Nance methodology themselves, we are working in collaboration with their experimental work to obtain images that are fed into the image processing pipeline by other researchers in the lab.

Of the four undergraduate students who continued data science research, the senior undergraduate student visualized and analyzed circularity in brain cells before graduating and is going to work in the chemical engineering industry. The senior with the circularity visualization project provided feedback that the program allowed them to explore whether a data science-focused or computational-focused industrial career was a good match for their personality. A junior interested in applying for chemical engineering PhD programs explored a brain cell morphology project centered around differences from location of injury and across brain regions by combining the methodology from TEXTILE with their own code for deeper analysis. The student provided feedback that the TEXTILE program helped them decide between pursuing a master's or PhD as their terminal degree. The sophomore student interested in applying to chemical engineering PhD programs is working on adapting a machine learning pipeline used by the lab for better analysis of individual cells. The sophomore joined the program with the original intention of applying for a PhD after completing their undergraduate degree and provided feedback that the program reinforced their interest in this goal. Finally, the freshman student was already participating in research in the lab in a wet-lab specific project but had minimal coding experience before the TEXTILE program. After the TEXTILE program, the student expressed interest in a computational research project that combined their wet lab expertise with increased computational development and application. The student was also encouraged by their learning during TEXTILE to participate in coding marathons and subsequently enrolled in additional computer science courses in their undergraduate curriculum and has expressed interested in pursuing a PhD in computer science.

Graduate Student Trainees

Of the graduate students, the masters' students both used the methodologies taught in TEXTILE directly in their masters' theses and further developed the methodology in the lab with addition of a software comparing common cell thresholding methods and morphological skeletonization of highly branched brain cells. Of the PhD students, both students who went through the TEXTILE program wrote a collaborative experimental paper evaluating oxygen-glucose deprivation in brain slices.[19] This paper represents a published outcome from the TEXTILE program, where experimental wet lab researchers with expertise in brain slice models and nanotherapeutic development integrated the expertise of the PhD students in the lab who developed TEXTILE to utilize and analyze cell feature analysis in their investigation of nanoparticle-cell interactions.

DISCUSSION

Student Feedback and Reflections

Fifty percent of students who participated in TEXTILE provided feedback through the Google Forms. However, most of the students gave verbal feedback during check-in portions of the main module, via personal email, or Slack Technologies, Inc. messages. Major and common feedback and suggested responses for addressing this feedback in the future have been compiled (Table 5).

Limitations and Areas for Improvement

The current limitations for TEXTILE fall into four main categories: experiment integration, accessibility, cohort size, and time to develop the program. In its current state, the modules developed for TEXTILE are mainly data science and image processing related. The heavy coding necessity of data science makes the modules easier to develop for an online classroom. We integrated the experimental aspects of the lab by using data and images produced by the lab. However, many students wished for more understanding and background of the experimental procedures that produced

TABLE 5
Student Feedback and Instructor Response

Student Feedback	Instructor Response
• Images too small in presentations, Jupyter Notebook too small of presenter	• Provide technology suggestions to encourage students to use larger screens • Enlarge text on all Jupyter Notebooks or set a reminder to Zoom in during main module sessions
• Coding errors are frustrating	• Encourage students that errors are a part of coding and do not relate to student capabilities • Provide examples of errors in instructors' own coding experiences
• Python in image processing modules still too advanced	• Provide open office hours for students to complete Intro to Python Learning exploration alongside an instructor
• Group discussion after pre-module activities was fun and a favorite part	• Build group discussion specific time into synchronous format • Encourage group discussion via Slack for asynchronous format
• Jupyter Notebook needs its own introduction before the modules begin	• Develop an Intro to Jupyter Notebook module to go alongside the existing Intro to Python Module
• Pacing and check-in with audience too much/ too little	• Assess coding experience of audience before each synchronous module • Provide instructions for utilizing Zoom emoji responses for slow down and speed up and encourage students to utilize these options

this data. To bridge the knowledge gap, we recommend adding in video lessons of wet-lab methodologies that result in the research data used within individual modules.

One of the major limitations of TEXTILE is accessibility; not all students have access to the same technology. Students accessed lessons using laptops, phones, and tablets that have varying screen sizes and computational power. Additionally, some students had access to better internet connections while others encountered difficulties connecting to Zoom lessons or streaming videos without buffering. To overcome accessibility issues, an instructor could apply for technology grants and provide students with equitable access to internet, such as prepaid Wi-Fi hotspots and technology for the duration of the program.

A third limitation of TEXTILE is the cohort size. Since group discussion and interaction between students in the program and researchers in the lab were considered valuable by the TEXTILE development team and feedback from the program, the size of cohorts must be maintained so the instructor can provide sufficient one-on-one interaction. We tested TEXTILE at a 15:1 student:teacher ratio and received feedback that the group was verging on too large for adequate assistance and discussion. However, additional cohorts will need to be taught to test the boundaries of the student:teacher ratio. Other ways to increase cohort size while still allowing students to receive adequate time with instructors include expanding the number of instructors and obtaining funding for staff to either complete lectures, handle administrative tasks, or manage the platform development and sustainability.

Finally, a limitation of TEXTILE is the amount of time it takes to develop a program. To identify applicable methodologies – those with a high number of features and complex, multi-layered experimental groups – and develop the initial curriculum is a significant time sink before the program is even created. Additionally, teaching the course annually or semi-annually is equivalent to a small-to-regular sized teaching load depending on the number of modules and size of the cohort. However, once the curriculum is developed, we believe the benefit of a sustainable and improvable laboratory training curriculum will save the lab time in the long run. Additionally, the modules can be pre-recorded and hosted online rather than in a "live" class to decrease time after initial start-up. There is flexibility in how any individual lab may adapt the TEXTILE curriculum to meet its needs.

In addition to overcoming the current limitations of TEXTILE, there are three other areas we identified for improvement. The first area of improvement is better assessment of students' current coding abilities. To improve student confidence with the modules, additional modules and office hours could be incorporated to encourage students to utilize already existing, open-source learning platforms for Python, Jupyter Notebook, and data science while developing additional introductory modules for confidence in research specific methodologies. To understand student coding experience, implementing pre-series, mid-series, and post-series student feedback on their current coding capabilities and confidence will allow the lecturer to tailor pace better to the current cohort. Additionally, deep exploration options can be added to each module so students can choose to take on a more advanced task when the pace may be slower than they desire.

Another area for improvement we identified was incorporating methods for students to write their own code – most of TEXTILE teaches students to run and modify existing methodologies. TEXTILE could build in individual projects related to lab research by identifying specific features or tasks that need to be developed. Additionally, incorporating code review sessions for students to discuss their code amongst each other and with the instructors running TEXTILE that currently utilize those specific methodologies could increase code literacy.

The final area of improvement we identified was larger impact of the program within the research lab's greater chemical engineering department. Since the TEXTILE modules train students to run data science-based experiments with recently published research, integrating the program with existing curricular labs or capstone programs could enhance undergraduate learning on a broader level while providing students hands-on connections between their classes and cutting-edge research. To increase the impact of the TEXTILE program, we propose that after research laboratories develop and test their own within-lab curriculum for training students, the research laboratories also team with curricula laboratories to alter the program for a broader educational purpose. Connecting research labs to curricula labs through the TEXTILE format could foster natural relationships between a university's research initiatives and undergraduate education while strengthening students' skills in data science and cutting-edge techniques.

Diversity, Equity, and Inclusion

Of the student population, 9 of the 15 students (60%) identify as female. As females in engineering and data science are underrepresented,[21] the opportunity for these students to learn in a majority female group represented an inclusive engineering training space. Other student demographics were not recorded or provided in publication due to the small size of the cohort. The program was free to all students, and technology accessibility issues are addressed in the limitations and areas for improvement section. To make the program equitable, we aim to apply for grant funding for technology allowances so students have equal access to technology including laptops and internet access in the future.

CONCLUSIONS

TEXTILE is an educational framework for developing modules and educational pathways to teach students research lab specific methodologies that utilize data science. We successfully taught 15 students – ranging from high school to doctoral students – Python-based image processing to segment cells from immunofluorescent images of brain slices. Many of these students also chose to continue in our research lab and are now working on independent or graduate student mentored research projects. In the future, TEXTILE will be expanded to incorporate machine learning based pipelines, experimental wet-lab methodology videos, and grant funding for student accessibility support. TEXTILE is a method for laboratories dedicated to student growth and training that decreases the amount of time spent training each student by developing a robust and interesting lab-specific curriculum, therefore increasing the research lab's capacity for student training and impact.

ACKNOWLEDGMENTS

The authors thank student participants in TEXTILE and their insightful feedback on the development of the TEXTILE program. The authors thank the Data Intensive Research Enabling Clean Technology (DIRECT) NSF National Research Traineeship (DGE-1633216) (H. Helmbrecht), National Institute of General Medical Sciences Grant #R35GM124677 (E. Nance), and NSF Harnessing Data Revolution Grant #OAC-1934292 (H. Helmbrecht, E. Nance) for funding support.

REFERENCES

1. Ashraf C, Joshi N, Beck DAC, and Pfaendtner J (2021) Data Science in Chemical Engineering: Applications to Molecular Science. *Annual Review of Chemical and Biomolecular Engineering*. 12(1):15-37. DOI: 10.1146/annurev-chembioeng-101220-102232.
2. Duever TA (2019) Data science in the chemical engineering curriculum. *Processes*. 7(11):830. DOI: 10.3390/pr7110830.
3. Amaral LAN and Ottino JM (2004) Complex systems and networks: Challenges and opportunities for chemical and biological engineers. *Chemical Engineering Science*. 59(8-9):1653-1666. DOI: 10.1016/j.ces.2004.01.043.
4. Pal S, Mondal S, Das G, Khatua S, and Ghosh Z (2020) Big data in biology: The hope and present-day challenges in it. *Gene Reports*. 21:100869. DOI: 10.1016/j.genrep.2020.100869.
5. Reisel JR, Walker CM, Cancado L, and Mitrayani D (2016) Successful undergraduate research experiences in engineering: Student, faculty, and industrial perspectives. *Proceedings ASEE Annual Conference*, available at https://monolith.asee.org/public/conferences/64/papers/14409/view
6. Madan CR and Teitge BD (2013) The benefits of undergraduate research: The student's perspective. *The Mentor: An Academic Advising Journal*. 15:1-3. DOI:
7. Beck DAC, Carothers JM, Subramanian VR, and Pfaendtner J (2016) Data science: Accelerating innovation and discovery in chemical engineering. *AIChE Journal*. 62(5):1402-1416. DOI: 10.1002/aic.15192.
8. Collins TW, Grineski SE, Shenberger J, Morales X, Morera OF, and Echegoyen LE (2017) Undergraduate research participation is associated with improved student outcomes at a Hispanic-serving institution. *Journal of College Student Development*. 58(4):583-600. DOI: 10.1353/csd.2017.0044.
9. Russell SH, Hancock MP, and McCullough J (2007) Benefits of undergraduate research experiences. *Science*. 316: 548-549. DOI: 10.1126/science.1140384
10. Fechheimer M, Webber K, and Kleiber PB (2011) How well do undergraduate research programs promote engagement and success of students? *CBE—Life Sciences Education*. 10(2):156-163. DOI: 10.1187/cbe.10-10-0130.
11. Zydney AL, Bennett JS, Shahid A, and Bauer KW (2002) Impact of undergraduate research experience in engineering. *Journal of Engineering Education*. 91(2):151-157. DOI: 10.1002/j.2168-9830.2002.tb00687.x.
12. Joshi M, Aikens ML, and Dolan EL (2019) Direct ties to a faculty mentor related to positive outcomes for undergraduate researchers. *BioScience*. 69(5):389-397. DOI: 10.1093/biosci/biz039.
13. Mcintyre N, Amelink C, and Bokor J (2020) Career development impacts of a research program on graduate student and postdoc mentors. *Proceedings ASEE Virtual Annual Conference*. DOI: 10.18260/1-2--34260.
14. Limeri LB, Asif MZ, and Dolan EL (2019) Volunteered or voluntold? The motivations and perceived outcomes of graduate and postdoctoral mentors of undergraduate researchers. *CBE—Life Sciences Education*. 18(2):ar13. DOI:
15. Park JJ, Choe NH, Schallert DL, and Forbis AK (2017) The chemical engineering research laboratory as context for graduate students' training: The role of lab structure and cultural climate in collaborative work. *Learning, Culture and Social Interaction*. 13:113-122. DOI: 10.1016/j.lcsi.2017.04.001.
16. Hall EE, Walkington H, Shanahan JO, Ackley E, and Stewart KA (2018) Mentor perspectives on the place of undergraduate research mentoring in academic identity and career development: An analysis of award winning mentors. *International Journal for Academic Development*. 23(1):15-27. DOI: 10.1080/1360144x.2017.1412972.
17. Morales DX, Grineski SE, and Collins TW (2017) Faculty motivation to mentor students through undergraduate research programs: A study of enabling and constraining factors. *Research in Higher Education*. 58(5):520-544. DOI: 10.1007/s11162-016-9435-x.
18. Morrison ME, Lom B, Buffalari D, Chase L, Fernandes JJ, McMurray MS, and Stavnezer AJ (2020) Integrating research into the undergraduate curriculum: 2. Scaffolding research skills and transitioning toward independent research. *Journal of Undergraduate Neuroscience Education*. 19(1):A64. DOI:
19. Joseph A, Liao R, Zhang M, Helmbrecht H, Mckenna M, Filteau JR, and Nance E (2020) Nanoparticle-microglial interaction in the ischemic brain is modulated by injury duration and treatment. *Bioengineering & Translational Medicine*. 5(3). DOI: 10.1002/btm2.10175.
20. Helmbrecht H and Nance E *textile*, GitHub, 2021. https://github.com/Nance-Lab/textile
21. *2021 Profiles of Engineering and Engineering Technology*. American Society for Engineering Education: Washington, D.C. ❐