Towards an Asynchronous Course-based

Undergraduate Research Experience (CURE) Framework:

A Pilot Case Study in Remote Genomics Research

by

Danielle O. Alarid

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved March 2023 by the
Graduate Supervisory Committee:

Melissa A. Wilson, Chair
Kenneth Buetow
Katelyn Cooper

ARIZONA STATE UNIVERSITY

May 2023

ABSTRACT

Course-based undergraduate research experiences (CUREs) are strategically designed to advance novel research and integrate future professionals into the scientific community by making relevant discoveries through iteration, communication, and collaboration. With Universities also expanding online undergraduate degree programs that incorporate students who are otherwise unable to attend college, there is a demand for online asynchronous courses to train online students in authentic research, thereby leading to a more skilled, diverse, and inclusive workforce. In this case-study, a pilot CURE leveraging the data-intensive field of genomics was presented as an inclusive opportunity for asynchronous, online students to increase their research experience without having to commit to in person or extra-curricular assignments. This online CURE was designed to investigate the effects of trimming software on high-throughput sequencing data when analyzing sex differential gene expression. Project-based objectives were developed to asynchronously teach (1) the biology behind the research, (2) the coding needed to conduct the research, and (3) professional development tools to communicate research findings. Course effectiveness was evaluated qualitatively and quantitatively using weekly, open-response progress reports and an assessment administered before and after term completion. This pilot study exhibited that students can be successful in remote research experiences that incorporate channels for communication, bespoke and accessible learning materials, and open-response reports to monitor challenges and coping strategies. In this iteration, remote students demonstrated improved learning outcomes and self-reported improved confidence as researchers. In addition, students gained more realistic expectations to self-assess computational research skill-levels and self-identified adaptive coping strategies that are transferrable to future research projects. Overall, this framework for an online asynchronous CURE effectively taught students computational skills to conduct genomics research in addition to professional skills to transition to and thrive in the workforce.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION TO COURSE-BASED UNDERGRADUATE RESEARCH EXPERIENCES (CURES)

1. DEFINITION OF A CURE

Undergraduate research experiences have been shown to increase student interest in STEM careers, gains in research skills, and the likelihood that students will pursue graduate degrees (Russell et al, 2007, Lopatto, 2004, Lopatto, 2007, Paalman, 2002). With many biology students anticipating applying for medical school, undergraduate research experiences allow students to meet research experience qualifications, thus increasing their chances of being accepted into medical school (Cooper), as well as exposing them to alternative career possibilities. One caveat is that coveted laboratory positions are often limited and competitive to enter, with even fewer opportunities afforded to remote students. In response to the need for more inclusive undergraduate research experiences (UREs), there are national calls to transform traditional introductory biology pathways by implementing evidence-based science courses (Woodin, Carter, and Fletcher 2010)(American Association for the Advancement of Science 2015). As a result, course-based undergraduate research experiences (CUREs) have been implemented at a growing-number of institutions as a way for students to meet undergraduate degree requirements while gaining skills to produce publishable research.

Students who participate in CUREs are familiarized with cultural norms of scientific research and are better prepared for independent experiences and graduate school (Bangera and Brownell, 2014, Bennett, 2020). CUREs emphasize novel scientific gains and deviate from traditional laboratory curricula that are often referred to as “cookbook,” or procedures with predefined results that are consistent over each iteration (Brownell and Kloser, 2015, Hekmat-Scafe, 2016). CUREs are often compared to inquiry instruction as both ask novel scientific questions, but the impact of inquiry-based instruction remains relevant only to the classroom as student learning outcomes. The overall goal of a CURE aims to generate new information of interest to the greater scientific community while simultaneously training students in translatable research skills and giving them experiences communicating scientific results to stakeholders and peers. CUREs are designed to be iterative experiences that encourage students to collaborate, troubleshoot, and discuss results. Driven by genuine scientific research in the instructing faculty's area of interest, no two CUREs are identical, and questions or aims for subsequent CUREs can be built from assessing the results of each iteration (Auchincloss, 2014).

In addition to leading the course-research topic, faculty have reported benefiting from developing CUREs because they are able to embed research within their instruction schedules, increase opportunities for teaching assistants, and help recruit students for their labs (Brownell and Kloser, 2015, Brownell et al. 2016). Students who complete CUREs have a foundation of background knowledge in the instructing faculty's area of interest and experience using the tools that helped them answer relevant research questions. Conversely, faculty have experience with the student and can make more informed decisions when accepting applicants for their own labs or referring them to colleagues. Students have also been acquainted with members of the research team who acted as Teaching Assistants (TAs) in the course and have an established rapport that lends to an easier integration into a primary research lab. Students who take CUREs and faculty who teach them both have increased opportunities to conduct publishable research. Established lab members may also gain experience leading teams, obtaining TA positions, and designing reproducible research. Even CURE "failures" are beneficial experiences that show students how original research does not always go according to plan and challenges them to think critically of how to improve experimental designs using the embedded curricula (Gin, 2018, Goodwin et al, 2021).

The literature shows that most CUREs have been limited to in person or hybrid experiences, but not implemented for asynchronous, online students, even though online students have shown the same amount of interest in conducting research while reporting a significantly lower awareness of off-campus opportunities (Faulconer et al, 2020). Asynchronous students are often limited by their ability to attend research labs on campus due to geographic location, work schedules, life obligations, among other responsibilities, and yet have demonstrated the same ability to make learning gains in remote biology courses (Paul and Jefferson, 2019). Additionally, remote students are more likely to be from underrepresented demographics including first generation students, LGBTQ+ gender identities, and non-traditional, adult learners (Bangera and Brownell, 2014). As such, limiting CURE experiences to on-campus students only adds another barrier for remote students to overcome. Online research experiences can broaden the amount of historically underrepresented student participation in STEM, increase their application to (and likely enrollment into) graduate programs, and open additional career opportunities, thereby making STEM education more equitable and the future workforce more diverse (Cooper, 2019, Merrell, 2022, Faulconer, 2020).

The COVID-19 pandemic forced some CUREs to shift to online formats, but the literature did not show many options or recommendations for making these CURES completely asynchronous (DeHaven, 2022). The data-intensive field of bioinformatics offers undergraduate students the opportunity to not only conduct novel research, but to do so on their schedule, at their desired location. Students who exhibit skills in computation and programming are more prepared to answer biological questions and benefit from the high demand of these skills in the STEM industry (Wilson-Sayers et al., Bennett, 2020, Merrell et al, 2022, Gao and Guo, 2023). CUREs focused on the -omics (e.g., transcriptomics, genomics, proteomics) can be curated to foster inclusive opportunities for remote students to think like, use the tools of, and communicate like a scientist in an asynchronous setting by working with novel research questions, real datasets, and computational software (Brownell et al., 2015).

The asynchronous CURE not only allows students from historically underrepresented backgrounds obtain research experiences, but also provides experience in potential challenges endured in authentic research and allows them to identify coping strategies for these challenges in an environment where mistakes and failures come without severe penalty. Coping refers to the way students respond to extrinsic stressors that could increase anxiety and affect mental health. Transitioning into computational research (Forrester, 2022), especially in an asynchronous setting, can be challenging for students, yet managed with open communication (McInnerney and

Roberts, 2004) and effective coping strategies. Coping strategies can be adaptive or maladaptive depending on whether they resolve stressors or prevent resolution respectively and are composed of different themes as outlined by Skinner et al. (Musgrove et al. 2021). Identifying student-reported barriers helps make informed modifications to future iterations of the course and may be insightful for remote CURE initiatives hereafter.

2. PILOT GENOMICS CURE

The purpose of this study was to leverage the remote capabilities of the data analysis techniques used in genomics for students to understand how to apply reproducible coding workflows to answer novel research questions. In this pilot, students investigated the effect of common trimming software parameters on sex differential gene expression in human placental tissue by programming in R. This novel question was derived organically by the instruction team by applying a pre-established sex differential expression vignette from original research to a novel batch of placental tissue data (Olney et al. 2022). Students applied the reproducible vignette (https://github.com/SexChrLab/Placenta_Sex_Diff) to preprocessed datasets trimmed to specific parameters of interest, then made pairwise and full comparisons of their results. Students were introduced to the manuscript writing process to present their data interpretations and analyses and conducted constructive peer reviews. Throughout the course learning materials, the biology of the placenta, computational processes for statistical analysis, and means of effective, scientific communication were highlighted.

To measure the learning gains students experienced throughout the novel, asynchronous course, we investigated student outcomes of a pre- and post-assessment developed by an instructional team from a single research lab. Using a mixed method approach, we assessed the effectiveness of the course on student research and pedagogical outcomes by asking the following three questions: 1) Can a remote CURE increase student ability to interpret and analyze data? 2) How does a remote CURE impact student comfort levels in computational research? 3) What self-reported coping strategies did students use to overcome asynchronous challenges?

Our objective is to elucidate on the development and assessment of CUREs that foster inclusive opportunities for underrepresented, remote students seeking research experience and to identify the challenges and strategies observed for future program considerations. In this case study, quantitative and qualitative data from 13 students representing one instance were described and evaluated. We hypothesized that authentic research experiences conducted remotely using bioinformatics workflows can improve students' ability to interpret and analyze data and increase student comfort levels in conducting computational research. Using the pre- and post-assessment, we found significant learning gains and monitored changes in reported comfort levels. A weekly progress report provided qualitative data on student challenges and the coping strategies they identified to persevere throughout the course. Within this project, we describe recommendations for designing (Chapter 2) and assessing the outcomes of an asynchronous CURE (Chapter 3). Based on these results, we include recommendations for future iterations (Chapter 4) and present a publicly available framework that can be replicated and applied to a wide array of research questions pertaining to gene expression analysis.

CHAPTER 2

DESIGNING A CURE FOR ONLINE, ASYNCHRONOUS STUDENTS

1. FUNDING, IRB APPROVAL, AND CONSENT

This study was funded through the National Science Foundation IUSE program (Award number 2044096 to MAW, KC, and SB), Arizona State University School of Life Sciences Online Undergraduate Research Scholars program (to MAW, KC, and KB) and was considered exempt pursuant to Federal Regulations 45CFR46 by the Institutional Review Board (IRB), as evaluated by Arizona State University IRB (STUDY00013025). Students consented to participating anonymously and completed an external survey on self-efficacy in computational research from interdepartmental collaborators within the institution.

2. DESIGNING AN INCLUSIVE REMOTE CURE

Because CUREs are customized by faculty research interests, they can be designed to fit quarter or semester terms depending on the course structure and scope of the project. A benefit of offering quarter-long, online CURE options is that, if paired with a prerequisite course, both can be taken over the span of a single semester. The prerequisite course introduces students to computational research then acts as a baseline of experience for the subsequent CURE, thus expanding enrollment to include students new to programming. Unlike extracurricular research, CUREs offer students credit towards their degrees and embed research into schedules instead of being added on top of rigorous course loads, making it more inclusive for students who are limited in availability. For pilot designs, we recommend capping the number of students to ensure adequate instructional support for troubleshooting code and assistance with analysis. Instruction teams are recommended to consist of faculty to guide the research and instruct the course, research leads who specialize in the research topic, and teaching assistants with online learning experience. This allows students to have different contacts to reach out to and adds robustness to the student's overall experience.

THE BACKWARD DESIGN

The backward design for developing research-driven CUREs (Cooper et al., 2017) starts by defining the research objectives, determining acceptable evidence, then constructing a course framework based around the background knowledge, techniques, and thought-processes required to answer a central research question. Questions should be novel and of interest to the scientific community and can be derived organically from other projects and datasets within the primary lab. The instructional team should approach course construction with the overarching goal of producing at least one journal publication on the outcome of the research question and potentially another on CURE student outcomes. As a result, students can have their work contribute to the resulting research article or to continue research with the primary lab's team.

Once the research question has been established, the experimental design process can be broken down into major steps that are assigned into weeks, with the first week dedicated to getting acquainted with the required technologies and the final two weeks for manuscript submissions and peer reviews. This begins the process of developing the main learning objectives, which are further refined into project-based objectives (submodules) that asynchronously teach (1) the biology behind the research, (2) the coding needed to conduct the research, and (3) professional development tools to communicate research findings. By splitting the research main objectives into these three, distinct aims, there are clear breaks between theoretical, applied, and professional knowledge content that may help reduce the cognitive load students experience (Caskerlu, 2020). Consequently, the evidence-based framework is reproducible for future remote CUREs on three levels: (1) "Professional development" submodules may be replicated over any research topic, (2) "Coding" submodules may be replicated for research pertaining to specific types of data processing and analysis, and (3) "Biology" modules may be replicated for research using the same types of data.

After the desired research and learning objectives have been identified, acceptable evidence in the form of assessments, progress reports, and weekly uploads measure student outcomes. The pre- and post-assessment quantifies specific learning outcomes overall and can be taken from an 'off-the-shelf' template (Shortlidge and Brownell, 2016) or custom developed by the instruction team. We suggest a form of acceptable evidence for achieving pedagogical goals is the progress report which uses open-response prompts for students to self-report weekly successes, challenges, and goals. The report offers qualitative insight for course calibration through student testimony and helps the instruction team adjust contemporaneously. Lastly, weekly assignment uploads are evidence for research and professional development milestones. Each week students complete coding templates for research objectives and then submit their custom results, which can be reproduced, modified, and applied in a wide array of research contexts. In the final weeks, uploads in the form of figure storyboards, outlines, draft manuscripts, and peer reviews offer students a preview of the processes of how to communicate interpretations and analyses when publishing.

LEARNING MATERIALS

After the assessments are determined, the learning materials to answer the assessment questions are developed and procured. Instruction teams can find publicly available videos that cover specific topics, and can custom develop recordings and presentations for the lecture content. Selecting a larger quantity of shorter videos (~3 minutes), reduces cognitive load and makes learning materials more reproducible by easily replacing videos that lose relevance in future iterations. In lieu of textbooks with impertinent information, relevant studies and publications can be used as reference materials for developing bespoke text and can be made available to students to read and interpret for themselves. Additionally, the vast amounts of published bioinformatics workflows and R code can be modified for the specific research contexts and to develop the coding templates for the weekly upload assignments. By incorporating published research into the design, learning materials, and goals of the experience, students have more opportunities for acquiring professional development skills.

CONSIDERING ACCESSIBILITY AND INCLUSION

To afford each student equal access to learning resources and opportunities for course success, accommodations for disabilities should be at the forefront of course delivery (Gin, 2022). In one institution’s survey, biology students with disabilities who experienced the transition from in-person to online learning during the COVID-19 pandemic reported that accommodations like flexible assignment deadlines and frequent instructor communications regarding assignments and due dates would have increased their success in remote science courses (Barber et al, 2021, Gin, 2022). Flexible deadlines encourage students to interact with instruction teams to communicate delays as is customary in research labs, and reciprocal accommodations may be made by the students for the instructors in likely instances of course troubleshooting. Additionally, offering multiple modes of communication like email, video conferencing, and instant messaging platforms encourages students to use their preferred method to receive and communicate information.

OVERCOMING CHALLENGES

One of the most consequential and challenging aspects of online education formats is driving communication and support in the asynchronous, virtual classroom (Faulconer, 2020). As mentioned, most online students seek the virtual format because it allows them to further their education while tending to life events and responsibilities such as work or family obligations, or to join specific programs while living virtually any distance from the campus. Consequently, students must learn to pace themselves within their own schedules to ensure they are completing all course requirements in a timely manner while often balancing other rigorous STEM courses. Self-guided learning in an unprecedented research-course format had the potential to act as a stressor for students as they learned to calibrate their own comprehension and application while meeting deadlines. Additionally, some options of course communication required visible interactions with instructors and peers that may have been anxiety-inducing for some students (Hilliard, 2019) and asynchronous communication alone has been reported to make some students feel even more isolated (McInnerney and Roberts, 2004). Students may also experience anxiety in R programming, but with continued exposure and support, that anxiety has been shown to decrease over the duration of the course (Forrester, 2022).

3. IMPLEMENTATION OF THE PILOT CURE

The research institution that offered the pilot CURE had a well-established online program, ran for a 7.5-week quarter session (Fall 2022-B), and was open to third- and fourth-year undergraduate students, in addition to graduate students who were enrolled as online learners. The instruction team consisted of one tenured professor and one research scientist that both specialize in computational research, in addition to one graduate student with experience in remote learning and R. The pilot CURE was capped at 20 students for initial enrollment, 15 total students completed the course, and the 13 who completed both the pre- and post-assessments were the subjects of this study. Students were based around the globe and had varying levels of programming exposure, with the prerequisite, fundamental course as the minimum experience. The prerequisite course introduced them to computational research in a Linux/Unix shell environment to learn, apply, and write code to process high-throughput sequencing data.

EXPERIMENTAL DESIGN OF THE PILOT GENOMICS CURE RESEARCH-PROJECT

The ability to process and analyze RNA-sequencing (RNAseq) data to quantify RNA transcripts and measure gene expression is becoming ubiquitous across the life sciences, yet one preprocessing step is applied with inconsistent parameters, if applied at all. Trimming software is applied to filter poor quality sequence reads identified by quality scores, or probabilities of bases being called accurately in the sequencing process, and has been found to impact expression analyses of *Drosophila melanogaster* neurons when applied aggressively (Williams et al., 2016). When it comes to the effects of these processes on sex chromosomes, however, the research is limited. This oversight, in conjunction with access to the dataset used in the placental sex differences study by Olney et al. (2022), influenced the question: what are the effects of trimming parameters on RNAseq data when analyzing for sex differential gene expression?

To identify the parameters of interest, a literature review was conducted to survey the most common parameters cited for the trimming software of interest, Bbduk. Although parameters and software were often explicitly stated, reasoning for each parameter choice was seldomly disclosed. By investigating the effect parameters have on downstream analyses, students develop thought processes and reasoning for their own methodologies and for determining credibility when reviewing studies. The parameter ‘trimq’ sets a threshold for minimum quality scores and filters out bases of reads that do not meet that standard in the resulting output files; the parameter ‘minlen’ sets a threshold for the minimum length of the trimmed reads and furthers filters out reads that do not meet the criteria specified (https://jgi.doe.gov/data-and-tools/software-tools/bbtools/bb-tools-user-guide/bbduk-guide/). We defined the top two parameters mentioned, ‘trimq’ and ‘minlen’, and the top three choices (trimq = 0, 10, 30, minlen = 10, 30, 75) for each of these parameters at the time of course implementation.

Students were exposed to data preprocessing pipelines and trimming in the prerequisite command line course, so early learning objectives on the design of the CURE acted as an extension and reiteration of previous learning outcomes. The course dataset came from Batch 1 of the Olney et al. (2022) study and consisted of 10 biological female (XX) placenta samples and 12 biological male (XY) placenta samples procured by Yale Biobank and RNA sequenced by Yale Sequencing Center for the Sex Chromosome Lab at Arizona State University. The instruction team ran the preprocessing including trimming steps on the placental RNAseq data to each parameter of interest and generated the gene-count datasets prior to course commencement. However, to ensure novel results, the instruction team did not run the differential expression pipeline or perform any analysis. In fact, simulated data was generated to test the coding templates the instruction team produced as learning materials. Because students were to investigate the effect of trimming software on downstream analyses, raw, or untrimmed reads acted as the comparison standard.

Each student ran the untrimmed control dataset and then was assigned specific trimming parameter datasets with one ‘trimq’ value and another ‘minlen’ value to investigate in teams of two that were paired by complementary skill levels (high/low) and assigned by faculty familiar with the cohort from the prerequisite course. Each of the nine teams were assigned different parameter combinations of interest and students were able to compare pairwise results as partners using bar plots, Venn diagrams, multidimensional scaling (MDS) plots, volcano plots, and other data visualization tools. After the pairwise comparisons were discussed, full comparisons were made with the entire class by using upset plots to draw conclusions and report them in a draft manuscript. In the tradition of credible research, students were to conduct peer reviews on two to three manuscripts to enhance collective understanding and collaboration.

RESEARCH-DRIVEN LEARNING OBJECTIVES

Modeled after authentic research practices, a central goal of the course design was reproducibility. Learning objectives and materials were intended as a framework for differential expression analysis, with the only modifications necessary per iteration requiring updates to the study-specific Biology sections in the first two modules. The core learning objectives were modeled after the general procedure for determining differential gene expression with each module elucidating on the biology, statistical analysis, and application of the corresponding step of the differential expression pipeline (Table 1).

Learning Objective	Biology	Coding	Prof. Dev.
1. Understanding the research project	Research aims, placental function, sex differences in gene expression, RNAseq	Navigating in the Unix shell and connecting to the RStudio server	Effective google searches and reading scientific articles
2. Gene expression analysis with R	DE experimental design, effects of parameters	Linear modeling and normalization	Searching for R packages, sections of a scientific article
3. Differential gene expression (DGE)	Normalization, linear modeling for DEA	Plots: MDS, box, violin, jitter, combination	Solving coding errors
4. Comparing two DGE lists	Overlap between two conditions	Venn diagrams, correlation	Reproducibility

5. Comparing full range (>2) of DGE lists	Overlap analysis for higher number of comparisons	Upset plots	Writing a scientific paper
6. Writing a research paper	Citations, figure legends, sections of paper	Methods section	Authorship, ethics of data sharing
7. Peer review	Contextualizing results, literature searches	Reviewing and commenting	Peer evaluations

Table 1. Fall 2022-B: Pilot Genomics CURE Learning objectives by module and submodule. The learning objectives by week (rows) that were based on the experimental design of the research question. Each week, students would complete submodule objectives (Biology, Coding, Professional Development) to gain experience fulfilling the main module objective. The yellow cells will need to be updated for every iteration, whereas the green can be replicated for any quarter-long research experience, and the blue and green cells can be replicated together to answer most quarter-long differential gene expression questions.

THE PRE-ASSESSMENT AND POST-ASSESSMENT

Students took an identical assessment before and after completing the pilot CURE with scores and answers hidden to remove any advantages for the post-assessment. The intent behind the pre-assessment was to provide a baseline of student subject knowledge and the post-assessment was implemented for evaluating student gains overall, by aim, and in personal comfort levels conducting computational research. The pre-assessment and post-assessment were developed among the instruction team and were representative of each of the learning aims with 25 total questions, seven of which were categorized as biology, seven as coding, four as professional development, and seven Likert surveys of student comfort levels in computational research. Students received one attempt to complete the pre-test, but would each receive full credit for the assignment regardless of score. Students received a single attempt to answer questions written in check-all-that-apply, matching and multiple-choice for Likert survey formats.

To ensure students were getting the answers correct for understanding specific concepts, the check-all-that-apply and matching questions had varying amounts of distractors and answers to be selected, ranging in five to six total options for most questions. To reduce anxiety levels, there was no time limit on test completion nor video proctoring (Gin et al., 2022) required, however, a test submission was required to move on to subsequent modules (Appendix B).

THE WEEKLY PROGRESS REPORT

To maintain an authentic research experience, the instruction team opted to develop an open-response summative assessment for students to submit weekly, designed in the fashion of the progress reports submitted in the lead faculty's lab. To reduce student grading anxiety and to treat the assignment as an authentic PI update, report grading was flexible and geared towards completion over correctness. The report prompted students to list their successes in learned and applied content, perceived and technical challenges, how they addressed those challenges, and to plan for progressing through subsequent modules (Table 2).

Category	Prompt
Accomplishments	• List novel findings • Concepts learned • Coding solutions • List successful communication with instructors and classmates

Challenges and how you addressed them	• List specific challenges for the week
	    ○ If you did not have challenges, describe your strategies/background used to make this a challenge-free week
	• List your approaches for addressing this challenge (and if it is still outstanding)
	    ○ If you did not have challenges, describe how you helped others address a challenge (via Slack, meetings, group discussion)

Goals for the next week	• List how current module's objectives relate to next module • List goals of the next module • List content you want to revisit, or you think would be valuable to learn more about

Table 2. Fall 2022-B: Pilot Genomics CURE Weekly Progress Report. The categories of the open-response Weekly Progress Report include successes, challenges, how those challenges

were addressed (coping strategies), and upcoming goals. Exploring student challenges and coping strategies in an open response format was essential for instructors to modify weekly meeting topics as needed. As mentioned, to maintain the integrity of the novel research, the instruction team did not run the full course pipeline, and the weekly progress reports became fundamental for keeping the pulse of class progress and to identify any major troubleshooting needed.

WEEKLY UPLOADS

The instruction team developed bespoke coding templates that increased in rigor for students to complete in R and then submit for grading each module. The coding packets modeled reproducible code in RMarkdown format that can be communicated and shared amongst the class and in publications. Coding packets were submitted in the weekly uploads section and graded for completion by the research scientist who led their development. For the major projects, students produced key materials to communicate with the scientific community: figure storyboards, outlines, and a final manuscript. Students were provided with a manuscript template used by the lead faculty's lab for publication and access to the draft of the resulting course publication. To enhance collaboration and understanding, students conducted peer reviews on 3 manuscripts and followed a rubric for constructive feedback (Table 3).

Direction	**Prompt**
For each review you are assigned, please comment on the following:	• Are there any points or sentences that are confusing to you? • Do you feel there are places where the author needs more detail in order to make their point? • Are the figures clearly labeled in the figure legend and clearly explained in the results? • Can you follow the logic the author used to address their topic? Can you suggest an order for figures/results that you think would be more logical or give a better flow? • Are there any statements that you feel need references? • Are there any typos, spelling and grammar mistakes, or phasing issues that you can point out to the author? • Are there any parts of the paper that you feel are particularly awesome? Go ahead and bring attention to that with a reason you think it's awesome.

Please make sure to keep the following in mind for language:	• Keep your comments positive and helpful; remember that we are all in this together • Do not suggest things that are out of scope or that you know would take too much time to do • Read the entire report once all the way through before making comments so you know whether something you think is missing is just in a different order • Try to be as specific as possible

Table 3. Fall 2022-B: Pilot Genomics CURE Peer Review Rubric. The customized peer review rubric outlining comments and considerations for evaluating peer manuscripts.

LEARNING MATERIALS ARE CUSTOM AND EVERGREEN

Once the learning objectives were established for the three aims of each module, the instruction team surveyed publicly available videos for relevant materials and created custom videos when necessary. To minimize cognitive overload and maintain subject relevancy, most videos were under five minutes long and all were fully transcribed by the instruction team to provide students a text alternative. Shorter videos can easily be switched out in future iterations if more updated or apposite information becomes available. For the text of each lesson, in lieu of expensive textbooks, the instruction team researched relevant findings and used the curated videos to write bespoke text identifying the relevancy and providing references. This ensured students received pertinent lessons and did not have to filter through extraneous materials to understand the research project. Materials the instruction team found helpful but did not use as references were listed at the end of each submodule in an “Additional Resources” section.

CONSIDERING ACCESSIBILITY AND INCLUSION

The through line of selecting all course materials and supplies was consistent accessibility for all students, regardless of socioeconomic skill status or technical skill level. Textbooks and software were at no cost to students as project-specific materials were authored by the instruction team and the general content was procured from publicly-available resources. A course introduction video accompanied with an introductory lab meeting explaining the course format acted as an orientation to initiate students to the research-course dynamic. When it came to the technology, students had the option to use the RStudio Server within the institution’s high performance computing cluster which reduces downloads and stores large datafiles, or to install the free, publicly-available software on local machines. Communication considerations were at the forefront of the course design and asynchronously leveraged direct messaging platforms coordinated video conferences with transcriptions and recordings available, and Canvas-integrated tools such as announcements and email.

OVERCOMING CHALLENGES THROUGH EFFECTIVE COMMUNICATION

Students were given four different options to meet with the instruction team in real time throughout each week, and the class was initially polled to establish the schedule for each video conference. Optional lab meetings covered student-reported challenge topics and were recorded, transcribed, and posted with an announcement each week. Writing hours gave students the opportunity to discuss weekly assignments or other assistance and shared research hours gave students the opportunity to troubleshoot code with the instruction team on a biweekly basis. Shared research hours were offered in the evening and gave students on different schedules the opportunity to meet with the instruction team and share their screen as they report errors or gaps in understanding. Frequent instructor postings to Slack allowed instructors to ask about challenges, point out common errors, and identify areas to increase support.

In addition to being accessible, the instruction team was also flexible with deadlines and uploaded report formats as the cluster presented issues when it came to knitting, or printing, the report. If students communicated delays with the instruction team, deadlines were extended without penalty, as seen with manuscript submissions and peer reviews. Some students especially struggled in their weekly uploads, and the instruction team was able to leave guiding comments for students to make revisions and resubmit for partial credit.

CHAPTER 3

ASSESSING AN ASYNCHRONOUS CURE

1. METHODS OF ASSESSMENT DURING THE CURE

Due to the unprecedented structure of the for-credit, remote research experience, assessments and modifications were considered throughout the duration of the course. The concurrent modifications were dependent upon the qualitative feedback students provided each week in the progress report assignment, by email, or in the course Slack channel. Each week the instruction team would take the organic responses and review the challenges reported to identify trends and monitor for any major coding issues. Students were also prompted to list the coping strategies they used to address those challenges, or, in cases of no reported challenges, the background knowledge or previous experience they leveraged to be successful. The instruction team would synthesize and quantify the individual responses into a semantics bank to identify the top reported challenges and successes. Instructors would meet independently each week to discuss the reported challenges and how to address them, with most of the solutions becoming the topic of the following week's synchronous lab meeting discussion. These dynamic lab meetings created a virtual flipped classroom that derived lesson plans from student feedback and collaboration to find effective solutions specific content or code of concern.

In cases of troubleshooting, the interactive communication structure allowed the instruction team to discover areas of the code that were incorrect or needed debugging and provided students with the iteration and thought processes needed to effectively troubleshoot. If there were major troubleshooting issues discussed in Shared Research Hours or in Slack, the research lead documented problem-solving processes from identifying the issue to finding and executing solutions. These documents were then distributed to students with announcements made in communication platforms and by email, in addition to presenting the information in the recorded and transcribed lab meeting.

2. METHODS OF ASSESSMENT AFTER THE CURE

In this case study, we conducted a mixed methods approach to describe and evaluate data from the test scores and progress reports of 13 remote students representing a pilot genomics CURE. A single course was offered at Arizona State University as a 400 and 500 level elective for online undergraduates and graduates pursuing degrees in biology. As recommended by Brownell et al. (2018), student self-efficacy and self-concept were evaluated by an external research team and published independently. Because of this and the small sample size of the pilot course, student demographics questions were excluded from the assessments and not considered in this analysis. Each question was derived from the bespoke text within the modules and the figures used to assess student interpretations came directly from publicly available studies introduced in the course (Appendix A, Table 2). Question 25 discussed student research experience and was discarded from the analysis as it was informative for the instruction team throughout the course but not relevant for the post-assessment.

QUANTITATIVE ANALYSES

Students completed the pre-assessment at the initiation of the course and received full credit for their submission, regardless of percentage correct. To maintain the integrity of the post-assessment, actual grades and responses were hidden from student view and results from both assessments could not be accessed, even after course completion. Only first attempts at the examination and students who completed both assessments were evaluated. Student pre-assessment and post-assessment results were exported from the Canvas learning management system (LMS) where they were administered and graded. The LMS only issued full credit for check-all-that-apply questions that had every selection correct, and deducted points for any distractor selection. For the purpose of this analysis, each question was worth 1.00-point total, and scores ranged depending on the number of options, correct answers, and distractors available per question. The eighteen total questions on biology, coding, and professional development learning outcomes were evaluated independently from the seven Likert survey questions. To investigate this question, we used the `t.test` function for a two-tailed, paired t-test to test for significance between pre-assessment and post-assessment results overall, by learning aim, and self-rated skill level. The t-test was efficacious in comparing for significant differences between the paired dataset from the same subjects at different time intervals and the two-tails account for scores that increase, decrease, or remain static. If the difference between pre-assessment and post-assessment scores was found to be statistically significant, we conducted a Cohen's D test using the `cohen.d` function in R to check for effect size and whether our results could be applied to the larger population. Student and question gains were evaluated by the magnitude of change across the varying scores by using the formula [(post-assessment - pre-assessment)/pre-assessment] to calculate normalized learning gains (NLGs) (Dehaven 2022, Paustian et al. 2017, Colt et al. 2011). Because the check-all-that-apply format allows for many different combinations for answer selection, questions of interest were further analyzed by frequency of selection of individual answers and distractors (incorrect options). Increases in correct answer options exhibit learning gains while decreases show learning gaps. Conversely, increases in distractor scores demonstrated misconceptions gained while decreases inferred learning gains.

The second research question investigated the change in student comfort levels before and after the course, and asked students to rate their personal comfort performing computational research competencies, such as operating Linux/Unix, online collaboration with peers, R programming, and reading and writing scientific manuscripts. One answer was selected from a multiple choice 5-point Likert scale from very uncomfortable to very comfortable (including a neutral option). Likert survey questions 19 - 25 (excluding Q21) were evaluated for change in comfort levels using the `likert` package in R. One question (Q21) prompted students to self-rate their skill level from novice to expert and was evaluated separately from the Likert questions.

Engagement in communication platforms was quantified by exporting individual analytics from the Slack channel. At the time of this study, there were no grading options between Canvas and Slack available. There were no requirements for Slack engagement, but students were strongly encouraged to, at minimum, read the channel for common troubleshooting issues. We used the ‘days active’ metric to evaluate how many days students were loading the channel page out of the total 51 days of the course duration which did not require any posts or replies. The ‘messages posted’ metric quantified the number of messages a student posted in any of the course-specific channels, as an initial post or reply. To sort students by engagement, or frequency of message posting, we found the average engagement by dividing the total conversation (100%) by the number of students (n = 13) and set a posting threshold for average or higher as high engagement, half of the average to the average as moderate engagement, and half of the average to no posts as low engagement. The Slack metrics were plotted against student NLGs and Pearson’s correlation coefficient was determined using the `cor` function in base R.

QUALITATIVE ANALYSIS

The weekly progress reports contained student testimonials on challenges experienced throughout the course and how those challenges were addressed, or the coping strategies used to complete the course. As the course was running, challenges were clustered into categories: biology and statistical analysis, coding, professional development, cognitive load, personal, and other. Including the absence of a coping theme, there are fifteen in total: five are adaptive, including problem solving, support seeking, information seeking, self-reliance, and cognitive restructuring. Accommodation, negotiation, and distraction can be adaptive or maladaptive depending on how the student responded to the stressor, and themes that are generally maladaptive include escape, rumination, helplessness, delegation, and opposition (Skinner et al).

3. RESULTS

The first research question investigated if students' ability to interpret and analyze data improved after completing the CURE based on their assessment scores. The paired, two-tailed t-test resulted in a significant improvement in overall scores ($t = 3.7$, $p = 0.003$) with the mean score increasing from by 11.64%, from 65.96% (SD = ±12.92) to 77.61% (SD = ±16.77, $p = 0.003$)(Figure 1).

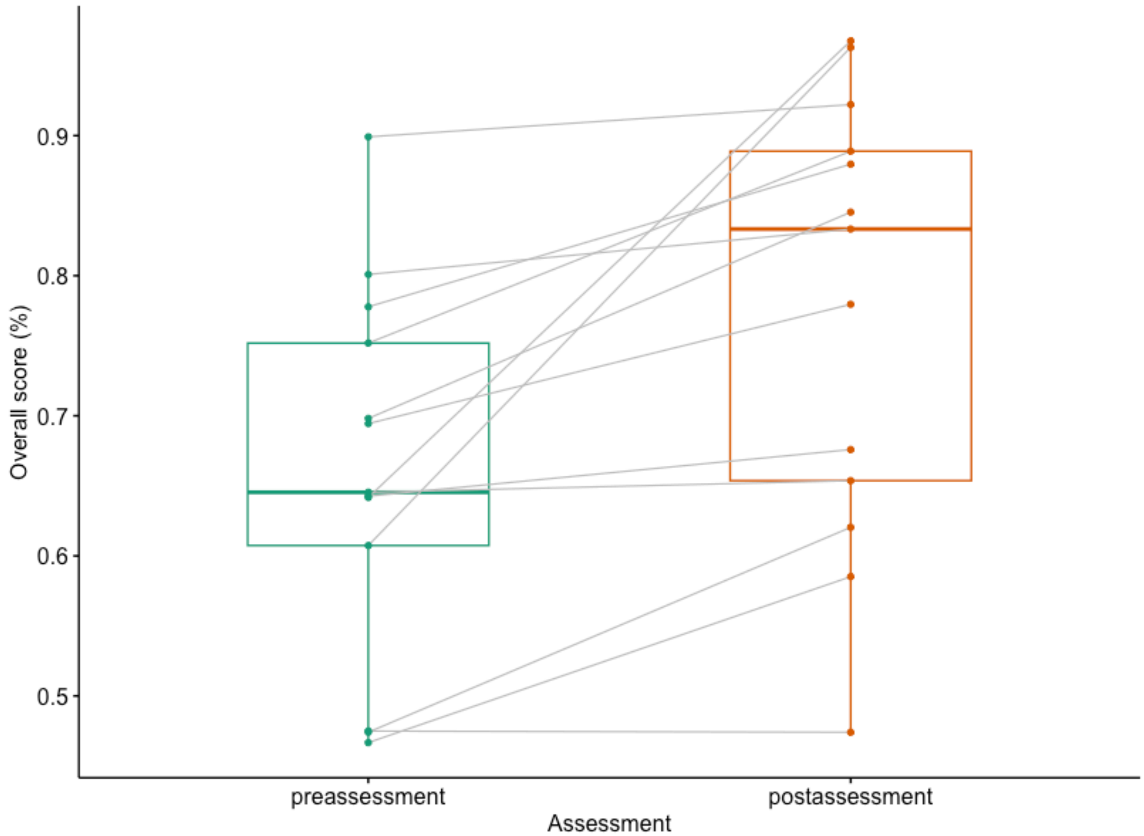


Figure 1. Student knowledge assessment before and after the course. Boxplots depicting mean student assessment scores before (green) and after (orange) completing the pilot GenomicsCURE. Each point represents one of the thirteen students who completed both the pre- and post- assessments and the lines connect to the same student's pre-assessment to post-assessment scores. The mean class score significantly increased by 11.64%, from 65.96% (SD = ±12.92) to 77.61% (SD = ±16.77, p = 0.003).

There were significant increases in average Biology (p = 0.0005) and Coding scores (p = 0.03). Students generally scored highly on Professional Development pre- and post-assessment scores were generally higher out of fewer questions and the increase in post-assessment scores was not considered significant (p = 0.10) (Figure 2).

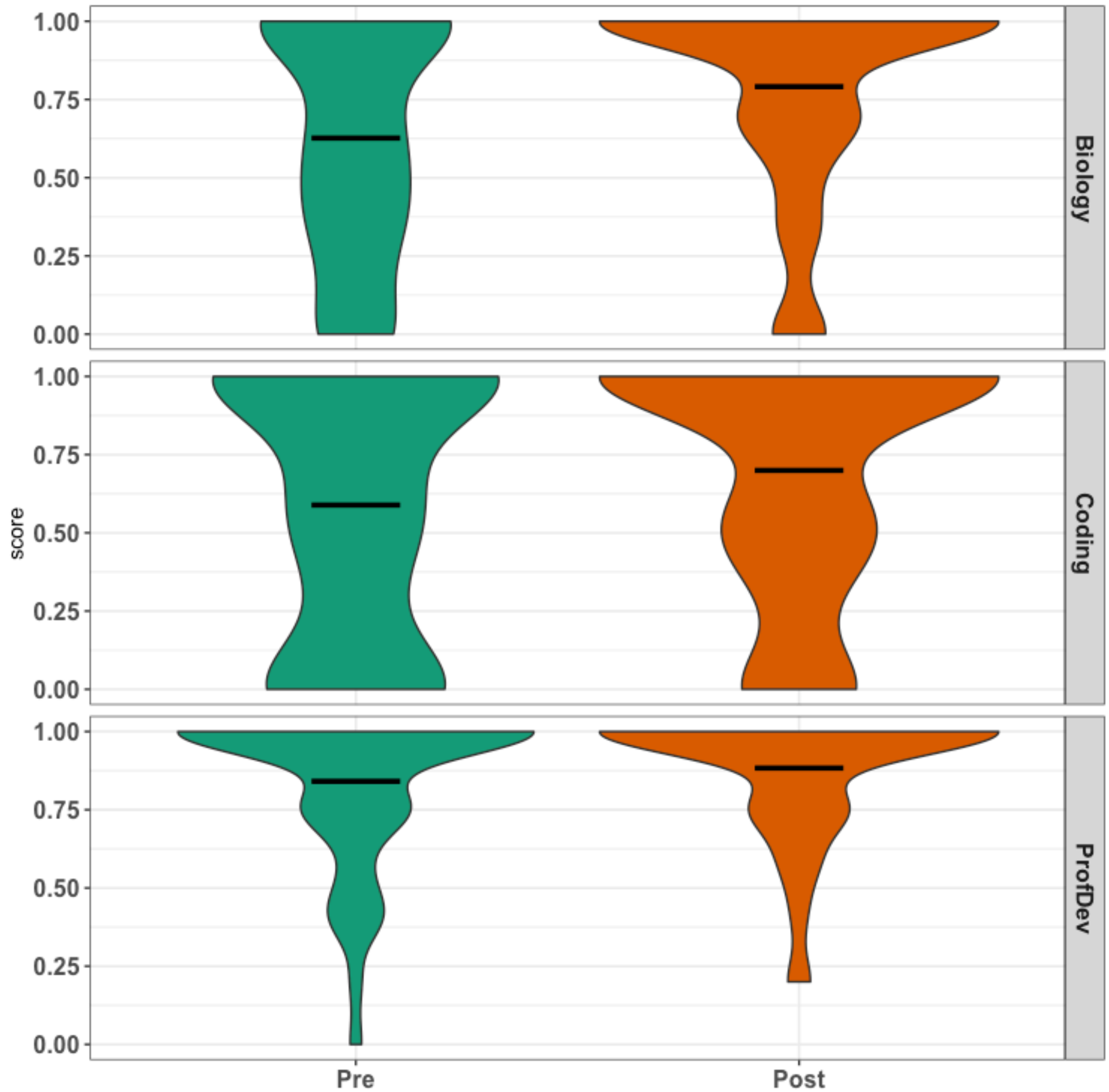


Figure 2. Range of pre- and post-scores for all questions divided by topic. Violin plots depicting each pre-assessment(green) and post-assessment (orange) scores for all questions divided by topic: (1) Biology (7 questions), (2) Coding (7 questions), and (3) Professional Development (4 questions). Varying thicknesses of each plot represent the distribution of scores and widths between each peak represent score densities, or clusters of score occurrences.

Mean scores of each question were generally higher in the post-assessment, with positive normalized learning gains (NLGs) in ~67% of questions (Q1, Q2, Q3, Q5, Q6, Q7, Q10, Q11, Q12, Q13, Q15, Q17), negative NLGs in ~11% (Q4, Q8), and no changes in ~22% (Q9, Q14, Q16, Q18) (figure 3).

Figure 3. (A) Average question pre- and post-assessment scores divided by topic and (B) difference in normalized learning gain. A. Barplot depicting mean pre-assessment (green) and post-assessment (orange) scores divided by submodule learning topics. Most questions exhibited gains except for Q4 and Q8 which had decreases in mean scores. Q14, Q16, and Q18 had no mean score changes between assessments. B. Barplot depicting the normalized learning gain (NLG) per question, categorized by topic: biology (green), coding (orange), and professional development (purple). NLGs represent the magnitude of average learning gain per question and

are determined by dividing the difference between post-assessment and pre-assessment scores by pre-assessment scores.

Q4 and Q8 experienced losses in learning gain and were analyzed for changes in frequency of answer and distractor selections between the pre-assessment and post-assessment to identify misconceptions gained or gaps in understanding. For example, in Q8, the question with the lowest mean NLG overall, 15.18% of students gained the learning outcome that saving R scripts with the data produced or analyzed allows for easier reproduction, modification, or sharing of protocols for future work (see figure 4). Differences between pre-assessment and post-assessment answer selections showed 7.68% less students considered R a language and environment for statistical computing and graphing, while 15.38% more students misconceived that R a language, but not a coding environment. An additional 23.08% of students misconceived that code in R was easy to read and did not need descriptive comments.

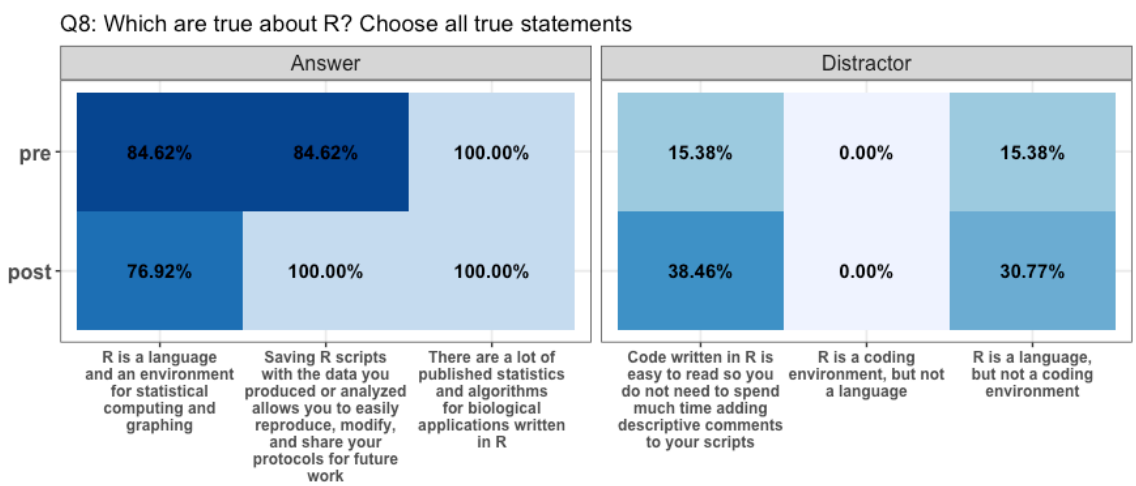


Figure 4. Student pre- and post-assessment answer selection frequencies for negative NLG questions. Heatmap evaluating frequency of answer choice selection for pre- and post-assessments classified by answers (correct) and distractors (incorrect) for Q8, which had the lowest difference (-10.26%) and NLG (-50.00%). Increases in the Answer category (left) show learning gains while decreases show learning gaps. Conversely, increases in distractor scores demonstrate misconceptions gained while decreases show learning gains.

The next research question investigated if comfort levels in computational skills changed after completing the CURE. According to the results from the Likert scale questions (Q19, 20, 22, 23, 24, 25), students no longer reported being very uncomfortable in any category overall, while the neutral option maintained or increased the original percentage reported across all categories (Figure 5). Combined comfort levels increased in online collaboration by 23% and decreased in operating Linux (-8%), R programming (-23%), and reading (-23%) and writing (-23%) scientific papers. The mean class Likert scores for the personal feelings section showed that the overall class remained within the same comfort level as reported in the preassessment, with reading scientific papers as the only level that decreased from comfortable to netural.

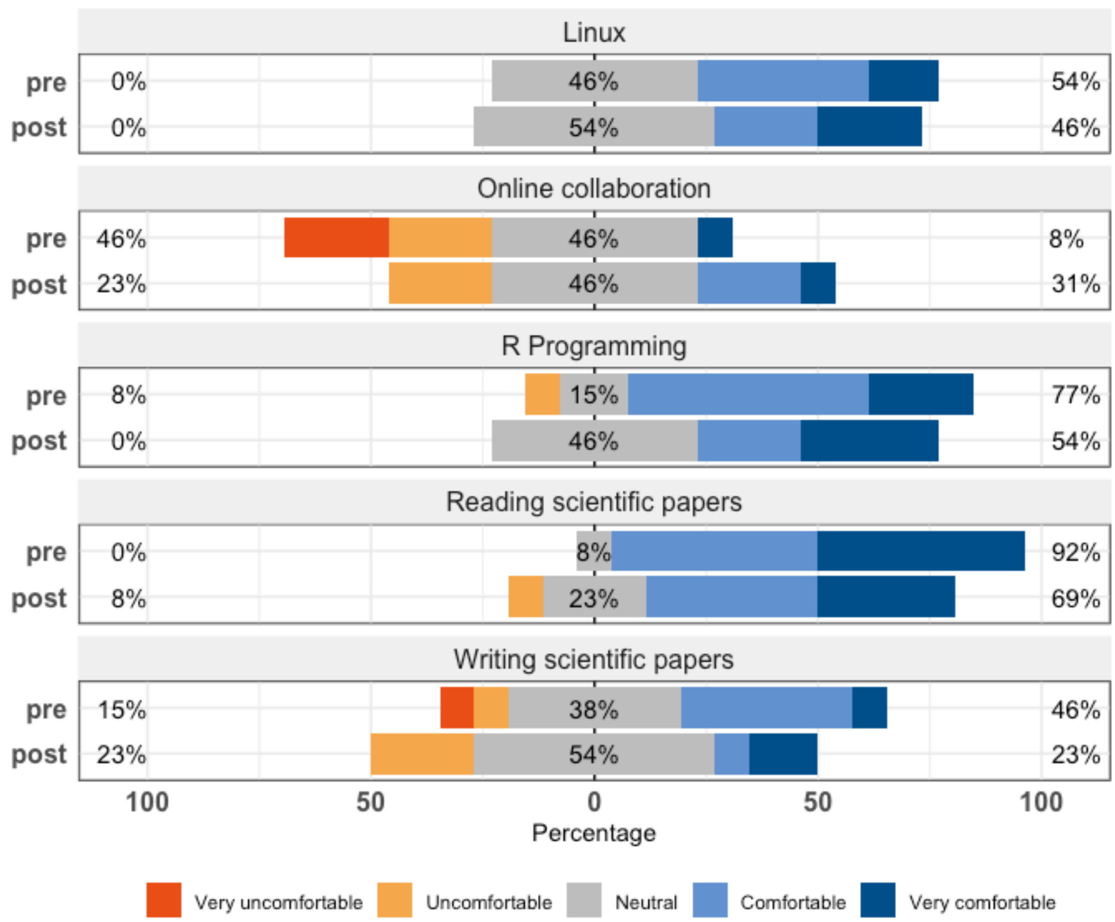


Figure 5. Self-reported comfort levels by question topic on the pre- and post-assessment. Red and orange values are indicative of discomfort, blue and navy values

represent areas of comfort, and gray is neutral. The percentages on the left combine discomfort scores, the percentages on the right combine comfort scores, and the percentages in the center are a direct representation of a neutral selection, differentiated by pre- and post-scores.

		Pre-assessment		**Post-assessment**			
Question	**Topic**	**Mean**	**SD**	**Mean**	**SD**	**+/-**	**P-value**
Q19	Linux/Cmd Line	3.692	0.7511	3.692	0.8549	0	1
Q20	Online collab	2.462	1.1266	3.154	0.8987	0.692	0.1
Q21	Expertise/SRSL	2.154	0.9871	2.538	1.1983	0.384	0.4
Q22	R Programming	3.923	0.8623	3.846	0.8987	-0.077	0.8
Q23	Reading papers	4.385	0.6504	3.923	0.9541	-0.462	0.2
Q24	Writing papers	3.308	1.0316	3.154	0.9871	-0.154	0.7

Table 4. Fall 2022-B: Pilot Genomics CURE Personal Feelings Assessment Scores. Questions and topics from the personal feelings section (columns 1, 2) of the course assessments. Pre-assessment and post-assessment class means and standard deviations are listed for each question, with the difference in scores shown in the (+/-) column. Between the two assessments, the scores for online collaboration and self-reported skill levels (Q20, Q21) increased while comfort in R Programming, reading papers, and writing papers decreased (Q22, Q23, Q24); command line (cmd) comfort remained unchanged (Q19). The p-values were determined using a paired, two-tail t-test that resulted in no significant differences ($p \leq 0.05$) in any of the personal questions.

For the question on self-rated skill level (Q21), the class pre-assessment mean was slightly over the “Advanced Beginner” rating with 2.15 (±0.99), and though the class post-assessment mean increased to 2.54 (±1.20), the increase was not significant ($p = 0.10$) (figure 6).

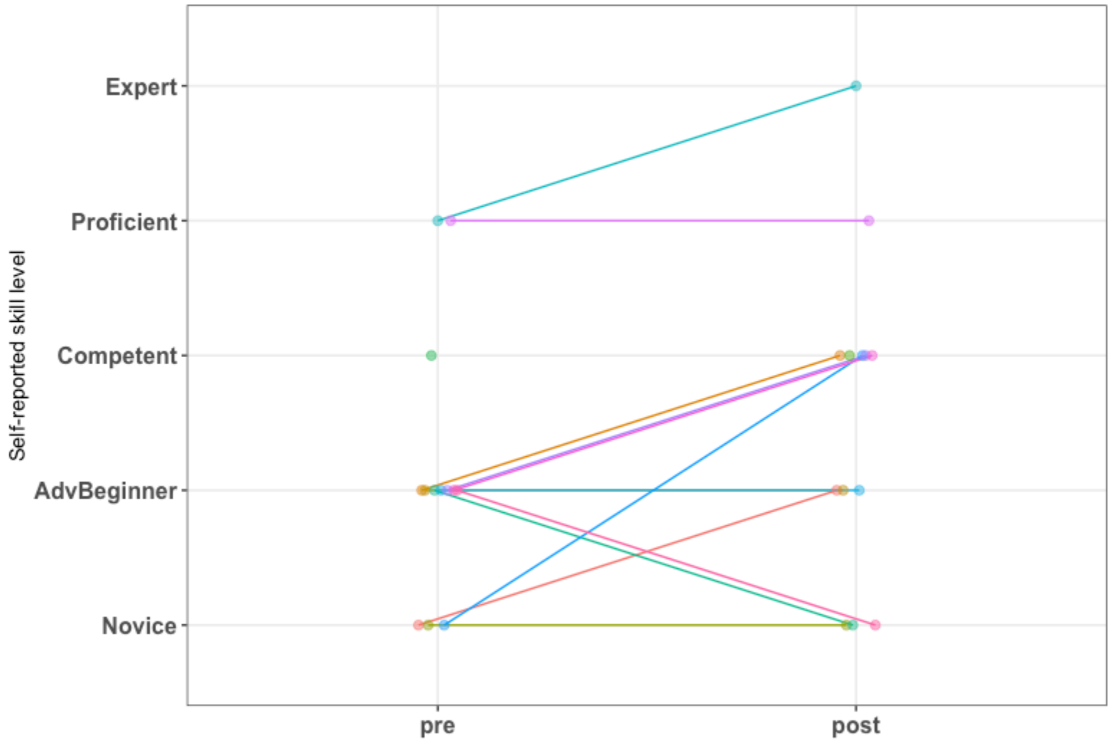


Figure 6. Student self-rated levels before and after the CURE. Student self-rated skill-level scores based on scale of increasing expertise, (1) Novice, (2) Advanced Beginner, (3) Competent, (4) Proficient, and (5) Expert from Q21 of the pre- and post- assessments. Each point corresponds to a different student and the line connects pre-scores to post-scores.

The last research question evaluated qualitative data from student progress reports to identify the challenges students experienced and the coping strategies students executed to address them. Overall, students reported the most challenges in Coding (56.91%), Personal obligations (12.77%), and Professional development (11.70%) themes (figure 7).

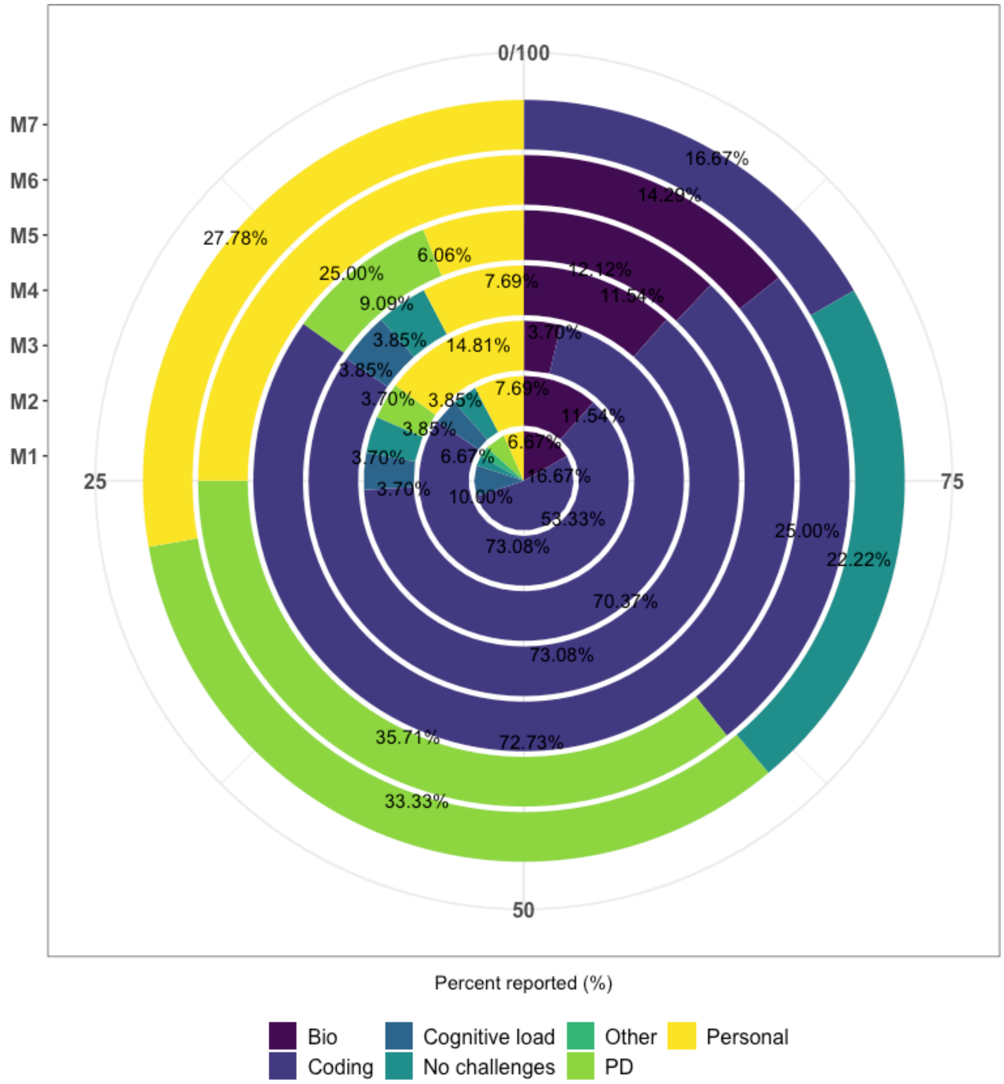


Figure 7. Self-reported challenges across each module. Pie chart of the challenges students self-reported in open-response progress report each module clustered by theme: Biology/statistical analysis (purple), Coding (indigo), Cognitive load (blue), No challenges (turquoise), Other (green), Professional development (light green), and Personal (yellow). The concentric rings represent each module, with the innermost ring representing the first module (M1) and subsequent modules radiating outward.

The coping strategies reported to address these challenges were mostly adaptive throughout the duration of the course, with maladaptive strategies occurring more frequently towards the penultimate module, or the week the mini manuscript was due. Overall, problem solving, support seeking, and negotiation were the top reported coping themes (figure 8).

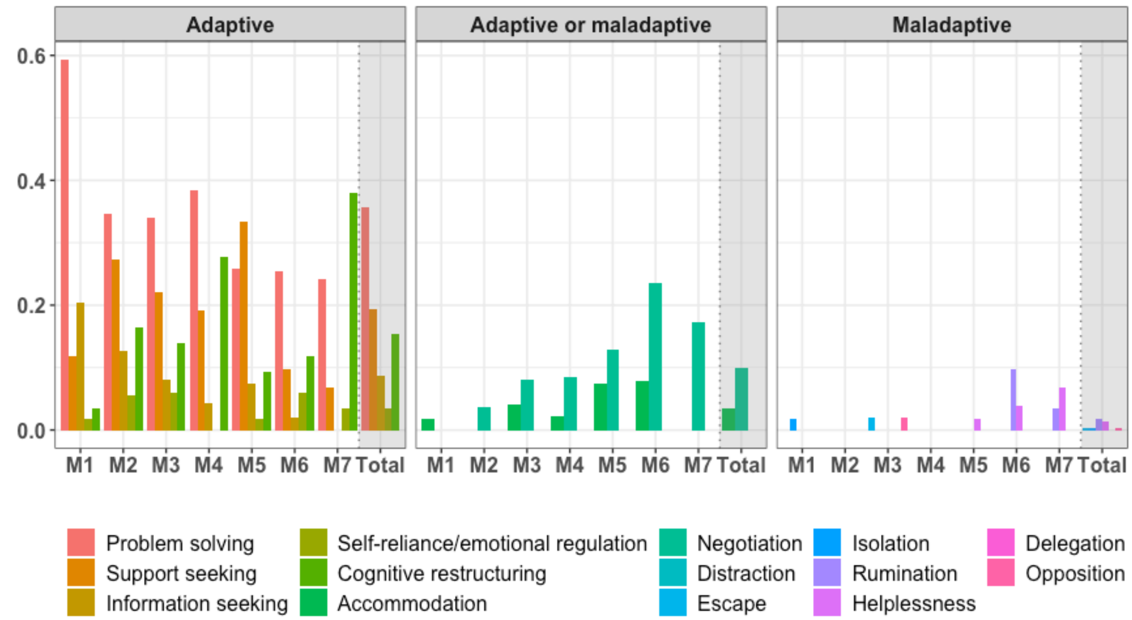


Figure 8. Coping themes categorized by coping type and frequency in each module. Coping strategies to overcome asynchronous course challenges self-reported in open-response progress reports each module, categorized as adaptive, maladaptive, or both depending on context. Adaptive themes include problem solving (red), support seeking (orange), information seeking (gold), self-reliance/emotional regulation (olive), and cognitive restructuring (green). Maladaptive themes include escape (light blue), isolation (blue), rumination (purple), helplessness (lilac), delegation (fuchsia), and opposition (pink).

Slack analytics of the course channel allowed us to quantify engagement and continue investigating a component of the support seeking theme for correlations with the mean class NLG (+28.63%). The mean percentage of days active was 50.23% (25.6 days) and there were 662 total messages posted by students across all course channels. The Pearson correlation between days active on Slack and NLG was weak and positive ($r = 0.3468$), whereas posting messages and NLG had a strong, positive correlation ($r = 0.8062$).

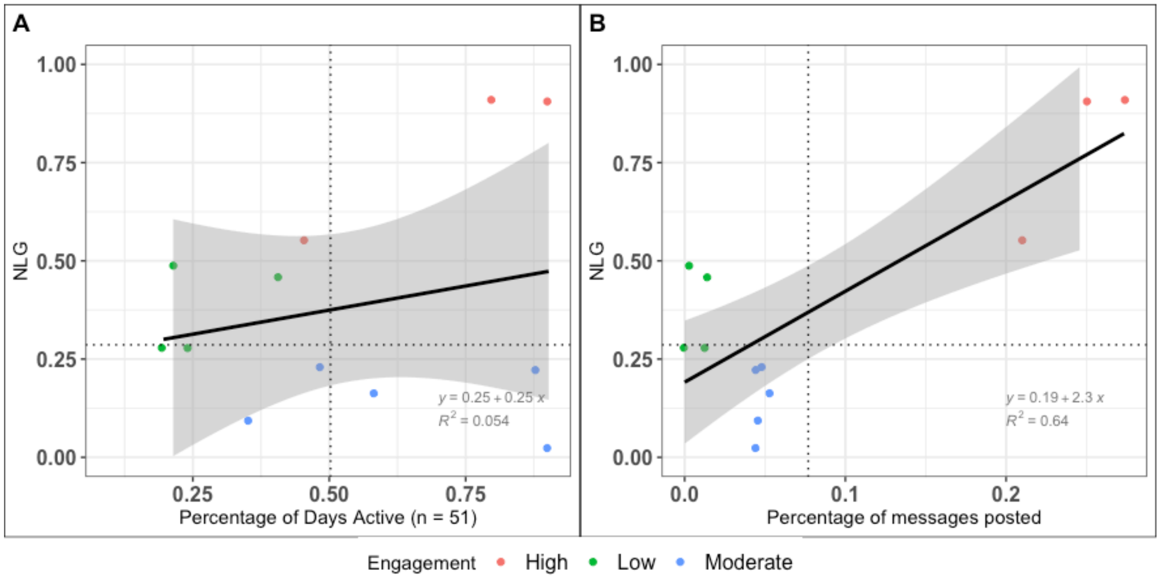


Figure 9. Normalized learning gains (NLG) by (A) Slack activity and (B) Slack messages posted. Student Slack engagement data was determined by pulling Slack analytics from course-specific channels from the day courses were made available until the end of the quarter. Engagement status was determined by dividing 100% of the conversation by the total number of students (n = 13 students) to determine if the theoretical average (7.69%) categorized as high (red, ≥ 7.69%), medium (blue, 3.81 - 7.68%), or low (green, ≤ 3.80). The horizontal, dotted line represents the mean class NLG (+28.63%). (A) Percentage of days active, or days of reading at least one channel or direct message, were determined by taking the individual's number of days active and dividing it by the total number of days in the course (n = 51 days). The resulting percentages were plotted against student NLGs to determine correlation. Vertical dotted lines represent the mean number of days active from Slack analytics (25.6 days or 50.23%) (B) Percentage of messages posted conveys the number of messages students sent in the course-specific channel (excluding direct messages) divided by total messages received across the channel (n = 662 total posts). Vertical dotted lines indicate the theoretical average for engagement status, 7.69%.

Ultimately, seven of the thirteen students sampled for this study (53.8%) expressed interest in joining a primary research lab after completing the asynchronous CURE. An additional two students expressed interest in authorship credit on the final course manuscript, resulting in a total 69.2% of students expressing interest in extending their research experience.

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS FOR ASYNCHRONOUS CURES

1. Conclusions from CURE course-assessment

We hypothesized that using the pre-assessment as a metric of pedagogical and research outcomes, overall, student ability to interpret and analyze data would increase after taking the remote genomics CURE. Initially, students received an average of 65.96%, but after taking the course, scores increased significantly to over 77.61%. Stratification of assessment results exhibited significant increases in biology and coding aims with students increasing their average scores per question in both categories. Scores for professional development sections also increased, however, these increases were not significant, most likely due to a high pre-assessment average. Scores for all but two questions increased, and students showed gains on topics pertaining to biology, statistical analysis, and interpretations of MDS plots and volcano plots.

Questions where NLGs decreased, Q4 and Q8, were examined by check-all-that-apply frequency of each option. In Figure 4, for example, the experience gained with R was integral to 100% of students selecting the option detailing the benefits of saving R scripts for reproducibility. The misconceptions students acquired in this question pertained to defining R as both a statistical language and coding environment. This may be due to students learning to distinguish R from RStudio by the separation of applications and reasoning became confounded between R as the statistical language and RStudio as the coding environment. An unexpected misconception came from the distractor stating code is easy to write and doesn't require descriptive comments. We believe this could be due to the ease of running the template code which was developed with accessibility in mind so students could complete within the quarter timeframe. The instruction team determined this was a poorly worded question and that it could have had answer options that were direct and not as open to interpretation. Further, students may experience question fatigue with the six options (Dillman, 2003), so reducing the options to five like the other questions may also contribute to improved performance in future iterations.

The second research question investigated the change in student self-reported comfort levels between the beginning and end of the genomics CURE, hypothesizing that student comfort levels would increase upon completion. We developed five questions that evaluated student comfort levels in Linux, online collaboration in forums (Slack), coding expertise in any language, R programming, and reading and writing scientific papers. Comfort in online collaboration and self-rated skill level (expertise) increased, but to our surprise, we found that comfort levels overall decreased in R programming and reading and writing papers, while Linux/Unix command line environments remained unchanged.

The increase in Slack comfort from advanced beginner to competent is conducive with the emphasis instructors put towards the best practice of reviewing for common troubleshooting solutions, at the minimum. This is particularly useful as over 100,000 companies use Slack as a communication tool (Slack, 2023) and it prepared students to effectively use tools that will be valued post-graduation. The decrease in R programming comfort levels was slight but may be due to the Dunning-Kruger effect in which due to their lack of skills in computational research, students were inexperienced in self-assessments and didn't realize how much of a learning curve they would experience (Zhou and Jenkins, 2020, Kruger and Dunning, 1999). Reading and writing papers disparities in comfort scores may be explained by the language used in the question itself. Instead of referring to "scientific papers", we postulate that directly stating "manuscripts" may render more accurate measures for both categories. Secondly, bioinformatics papers are structured differently than the general biology studies students may be more accustomed to, and this may have also contributed to decreased scores. Based on these results, we concluded that students gained more realistic expectations and an improved ability to self-assess computational research skill-levels.

Regarding the unchanged scores of 'neutral' in the Linux and command line category, tasks exhibiting these skills were limited, iterative of their prerequisite course experiences, and did not introduce new information. Because of this consistency, comfort scores may not have changed measurable amounts. Lastly, despite the decrease in overall comfort in several categories, student self-rated skill level (Q21 - Expertise) increased, albeit not significantly, and may demonstrate that although students became more aware of their discomfort in performing certain computational research tasks, they reported their skill levels continuing to advance. This suggests that students increased in confidence as computational researchers and, in agreement with the Dunning Krueger effect, became aware that identifying their own gaps in understanding was essential to building their skillset, as evidenced by this quote from the final weekly progress report:

> "I honestly can't believe that 7 weeks has pasted [sic] so fast. This has been an amazing experience and I am so lucky to have gotten to be a part of it. I honestly did learn a lot even if it may not have seemed like it. As an online student it's been really special to get to be a part of this experience. It just affirms my dream of going forward to graduate school."

In the last research question, we explored the weekly progress reports to identify qualitative data that informed on the challenges and coping strategies students experienced throughout the pilot remote CURE. We found that challenges in biology and statistical analysis decreased, and this is consistent with the steady decrease in the biology submodule materials as courses progressed and students generated more professional development documents, such as figure storyboards, manuscript outlines, manuscripts, and peer reviews. Each week, coding objectives were reported as the most challenging, with RStudio Server in the institutional cluster posing more system issues than expected. Within the progress reports, students openly expressed their concerns pertaining to coding and their visions for where they estimate they should be, as seen in one student entry where they stated they "still feel uninitiated in writing R code and remembering common commands to feel comfortable making changes in original Rmd script."

Because of this and the desire to continue computational research after course-completion, several students reported downloading R and RStudio software to their local machines, but an official query was not conducted. Finally, as the course progressed, the number of personal challenges reported each week increased from 6.67% in the first module to 27.78% in the final module. This was observed by the number of extensions that were requested and late submissions on the final two weekly upload assignments, the draft manuscript and peer reviews. Students were able to express accountability for personal delays and would exhibit several coping themes to assist in completing required assignments:

> "I waited until the last minute to complete this assignment, so I won't be able to get the help I need before submitting, but I plan to reach out to the instructors for help with determining findings from the charts in the code. I plan to do this not necessarily to help my grade, but because I genuinely want to get better at this so I can apply this skill later on."

After identifying this trend in increasing reported personal challenges, we sorted the coping strategies students identified into coping themes that were categorized by propensity to either solve problems (adaptive) or to prevent solutions (maladaptive). After taking each student's organic response and pulling out coping strategies, we found that in the first week, students heavily relied on problem solving (adaptive) to navigate and complete the initial module. As students became familiar with the course structure and expectations, they depended more on support seeking strategies such as Slack engagement, lab meetings, or communicating with instructors. Conversely, as the course progressed, there was an increase in strategies that could be either adaptive or maladaptive, such a negotiation, in addition to themes that were distinctly maladaptive. In the second half of the quarter, negotiation (adaptive or maladaptive) increased, typically in the form of asking for extensions and disclosing conflicting academic or personal commitments. Maladaptive strategies were minimal, except for cases of rumination or helplessness being disclosed in the final weeks. This corresponds with the increase in personal challenges observed in Figure 7 and the implicit need to foster inclusive learning environments for historically underrepresentation of remote students.

In the final analysis, we explored the increase in student-reported reliance and comfort in utilizing Slack as an educational communication platform. We found that simply loading and reviewing the channel had a weak positive correlation with increased learning gains, but those who were highly engaged in posting and loading, were more likely to experience above-average learning gains. Consequently, those who were moderately to highly engaged were more likely to express interest in continuing their research experience in a primary lab:

> "[I want to] [c]ontinue to learn the whole research process and hopefully allow more opportunities to continue to do these type of classes as my goal for my degree is getting into biological data research."

Even those who did not pursue extra-curricular research opportunities expressed appreciation for the course objectives as they relate to interdisciplinary research: "as I pursue my career in microbiology, I intend to bring my newfound knowledge of bioinformatics to the forefront. I desire to continue furthering my understanding in this field because I see a connection between these two fields of study."

2. DISCUSSION AND FUTURE DIRECTIONS

Undergraduate research experiences have been shown to increase student participation in STEM, retention in STEM degrees, applications to graduate or medical school, and overall employability. Course-based undergraduate research experiences (CUREs) generate curriculum directed towards developing publishable research while embedding these benefits within student schedules, and are advantageous for students, faculty, and institutional labs. There is a gap, however, when it comes to providing these evidence-based experiences for remote students who are historically underrepresented. This remote CURE was effective using the backward design to create research-driven learning objectives with aims in biology, coding, and professional development skills on a pilot sample size of 13 students. Leveraging the remote capabilities of genomics research, students significantly increased their ability to interpret and analyze data and increased in confidence in their skill-levels despite decreasing in comfort in areas of computation from initial perceptions. Throughout the challenges students experienced, they exhibited adaptive coping strategies such as problem solving, seeking support, and negotiation amongst commitments. Ultimately, students who engaged the most in the available communication channels exhibited higher learning gains and were more likely to continue in the lead faculty's primary lab.

There are limitations regarding the small sample size of the pilot CURE, as student factors such as demographics and full-time status, could not be addressed in analysis. Collecting data from future iterations and conducting a large-scale study, however, may offer additional insight. Based on the qualitative and quantitative data from the remote pilot, we recommended the several modifications for future asynchronous CUREs. Requesting CURE participant consent to access outcomes from the prerequisite course may help identify trends or challenges students experience and test if success in the first course is correlated to success in the CURE. When it comes to the assessments, monitoring student interest in applying for graduate school before and after course completion measures any increases as a result of CURE participation. Distinct and discernible language in assessments is necessary for accurate evaluation of student learning gains, and it is recommended to recruit question reviewers such as undergraduates, post-docs, and tenured faculty with different perspectives. Further, ranking each assessment question by Bloom's taxonomy (Freeman, 2011) may help in adjusting questions for future iterations.

In response to student and instructor feedback, we postulate that minor modifications to Canvas modules may increase student learning outcomes. In the pilot CURE, we introduced students to programming in R and RStudio in the second week but recommended initiating students in the first week using a publicly available R tutorial to increase experience generating code. To ease the challenges students faced submitting the draft manuscript in the final weeks of the course, we recommended incorporating weekly submissions of student-generated figures and analyses. Another consideration would be to adjust the progress report submission formats from PDFs to open-response quizzes for easier grading and qualitative data export. The instruction team agreed that adding a formative assessment such as a quiz, may help increase student learning gains. In respect to the learning materials, recorded lab meetings focused on challenging topics may be embedded into future modules and replace videos that become irrelevant. Considerations for scaling up to high-enrollment include automating open response grading by exploring software that can use semantic banks from student responses as baselines for processing open-responses and leveraging multiple graders for the weekly uploads and progress reports. Lastly, weekly Slack polls can quickly gauge student experiences and increase student interaction.

This pilot study exhibited that students can be successful in remote research experiences that incorporate channels for communication, bespoke and accessible learning materials, and open-response reports to monitor challenges and coping strategies. We found the pilot CURE not only improved remote student learning outcomes, but also improved student reported confidence as researchers. The pilot CURE helped remote students gain more realistic expectations and improved ability to self-assess computational research skill-levels. Students also self-identified adaptive coping strategies throughout the course that are transferrable to future research projects. The present framework trained students on the biology and coding behind differential gene expression analysis and may be reproduced for a wide array of CURE topics relating to genomics.

REFERENCES

American Association for the Advancement of Science. 2015. “Vision and Change in Undergraduate Biology Education: Chronicling Change, Inspiring the Future. A Report of the American Association for the Advancement of Science, Washington, DC.[Online.] H Ttp.” *Visionandchange.org*.

Auchincloss, Lisa Corwin, et al. "Assessment of course-based undergraduate research experiences: a meeting report." (2014): 29-40.

Bangera, Gita, and Sara E. Brownell. "Course-based undergraduate research experiences can make scientific research more inclusive." CBE—Life Sciences Education 13.4 (2014): 602-606.

Barber, Paul H., et al. "Disparities in remote learning faced by first-generation and underrepresented minority students during COVID-19: insights and opportunities from a remote research experience." Journal of microbiology & biology education 22.1 (2021): ev22i1-2457.

Bennett, Jennifer A. "The CURE for the Typical Bioinformatics Classroom." Frontiers in Microbiology 11 (2020): 1728.

Brownell, Sara E., et al. "A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data." CBE—Life Sciences Education 14.2 (2015): ar21.

Brownell, Sara E., and Matthew J. Kloser. "Toward a conceptual framework for measuring the effectiveness of course-based undergraduate research experiences in undergraduate biology." Studies in Higher Education 40.3 (2015): 525-544.

Caskurlu, Secil, et al. "Cognitive load and online course quality: Insights from instructional designers in a higher education context." British Journal of Educational Technology 52.2 (2021): 584-605.

Colt, Henri G., et al. "Measuring learning gain during a one-day introductory bronchoscopy course." Surgical endoscopy 25 (2011): 207-216.

Cooper, Katelyn M., Logan E. Gin, and Sara E. Brownell. "Diagnosing differences in what Introductory Biology students in a fully online and an in-person biology degree program know and do regarding medical school admission." Advances in Physiology Education 43.2 (2019): 221-232.

Cooper, Katelyn M., Paula AG Soneral, and Sara E. Brownell. "Define your goals before you design a CURE: a call to use backward design in planning course-based undergraduate research experiences." Journal of microbiology & biology education 18.2 (2017): 18-2.

DeHaven, Brian, et al. "Bootleg Biology: a Semester-Long CURE Using Wild Yeast to Brew Beer." Journal of Microbiology & Biology Education 23.3 (2022): e00336-21.

Dillman, Don A., et al. "Multiple answer questions in self-administered surveys: The use of check-all-that-apply and forced-choice question formats." Annual Meeting of the American Statistical Association. San Francisco, CA. 2003.

Faulconer, Emily K., et al. "Perspectives on undergraduate research mentorship: A comparative analysis between online and traditional faculty." Online Journal of Distance Learning Administration 23.2 (2020): 1.

Faulconer, Emily K., Charlotte Bolch, and Beverly Wood. "Cognitive load in asynchronous discussions of an online undergraduate STEM course." Journal of Research in Innovative Teaching & Learning ahead-of-print (2022).

Faulconer, Emily, et al. "Is a Framework of Support Enough? Undergraduate Research for Online STEM Students." Journal of College Science Teaching 51.3 (2022): 3.

Forrester, Chiara, et al. "Undergraduate R Programming Anxiety in Ecology: Persistent Gender Gaps and Coping Strategies." CBE—Life Sciences Education 21.2 (2022): ar29.

Freeman, Scott, David Haak, and Mary Pat Wenderoth. "Increased course structure improves performance in introductory biology." CBE—Life Sciences Education 10.2 (2011): 175-186.

Gao, Lei, and Miao Guo. "A course-based undergraduate research experience for bioinformatics education in undergraduate students." Biochemistry and Molecular Biology Education (2023).

Gin, Logan E., et al. "New online accommodations are not enough: the mismatch between student needs and supports given for students with disabilities during the COVID-19 pandemic." Journal of Microbiology & Biology Education 23.1 (2022): e00280-21.

Gin, Logan E., et al. "Students who fail to achieve predefined research goals may still experience many positive outcomes as a result of CURE participation." CBE—Life Sciences Education 17.4 (2018): ar57.

Goodwin, Emma C., et al. "Is this science? Students' experiences of failure make a research-based course feel authentic." CBE—Life Sciences Education 20.1 (2021): ar10.

Hekmat-Scafe, Daria S., et al. "Using yeast to determine the functional consequences of mutations in the human p53 tumor suppressor gene: An introductory course-based undergraduate research experience in molecular and cell biology." Biochemistry and Molecular Biology Education 45.2 (2017): 161-178.

Kruger, Justin, and David Dunning. "Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments." Journal of personality and social psychology 77.6 (1999): 1121.

Lopatto, David. "Survey of undergraduate research experiences (SURE): First findings." Cell biology education 3.4 (2004): 270-277.

Lopatto, David. "Undergraduate research experiences support science career decisions and active learning." CBE—Life Sciences Education 6.4 (2007): 297-306.

McInnerney, Joanne M., and Tim S. Roberts. "Online learning: Social interaction and the creation of a sense of community." Journal of Educational Technology & Society 7.3 (2004): 73-81.

Merrell, Laura K., et al. "Developing an Assessment of a Course-Based Undergraduate Research Experience (CURE)." Research & Practice in Assessment 17.1 (2022).

Musgrove, Miranda M. Chen, et al. "To cope or not to cope? Characterizing biology graduate teaching assistant (GTA) coping with teaching and research anxieties." CBE—Life Sciences Education 20.4 (2021): ar56.

Olney, Kimberly C., Seema B. Plaisier, Tanya N. Phung, Michelle Silasi, Lauren Perley, Jane O'Bryan, Lucia Ramirez, Harvey J. Kliman, and Melissa A. Wilson. 2022. "Sex Differences in Early and Term Placenta Are Conserved in Adult Tissues." *Biology of Sex Differences* 13 (1): 74.

Paalman, Mark H. "Undergraduate research, education and the future of science." The Anatomical Record: An Official Publication of the American Association of Anatomists 269.1 (2002): 1-2.

Paul, Jasmine, and Felicia Jefferson. "A comparative analysis of student performance in an online vs. face-to-face environmental science course from 2009 to 2016." Frontiers in Computer Science 1 (2019): 7.

Paustian, Timothy D., et al. "Development, validation, and application of the microbiology concept inventory." Journal of microbiology & biology education 18.3 (2017): 18-3.

Russell, Susan H., Mary P. Hancock, and James McCullough. "Benefits of undergraduate research experiences." Science 316.5824 (2007): 548-549.

Sayres, Melissa A. Wilson, et al. "Bioinformatics core competencies for undergraduate life sciences education." PloS one 13.6 (2018): e0196878.

Shortlidge, Erin E., and Sara E. Brownell. "How to assess your CURE: a practical guide for instructors of course-based undergraduate research experiences." Journal of microbiology & biology education 17.3 (2016): 399-408.

Shortlidge, Erin E., Gita Bangera, and Sara E. Brownell. "Faculty perspectives on developing and teaching course-based undergraduate research experiences." BioScience 66.1 (2016): 54-62.

Slack. “Why Nearly 80% of Fortune 100 Companies Rely on Slack Connect to Build Their Digital HQ.” Slack, https://slack.com/blog/transformation/fortune-100-rely-slack-connect-build-digital-hq.

Williams, Claire R., et al. "Trimming of sequence reads alters RNA-Seq gene expression estimates." BMC bioinformatics 17.1 (2016): 1-13.

Woodin, Terry, V. Celeste Carter, and Linnea Fletcher. 2010. “Vision and Change in Biology Undergraduate Education, a Call for Action--Initial Responses.” *CBE Life Sciences Education* 9 (2): 71–73.

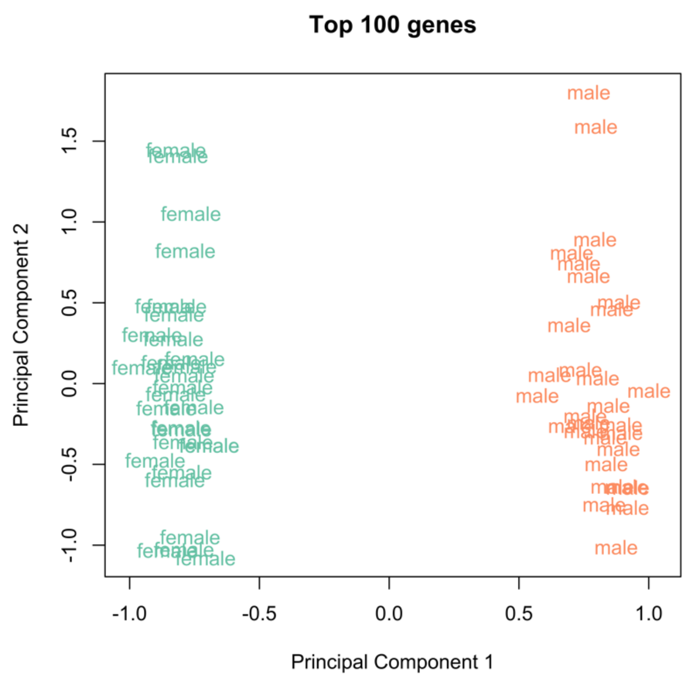
Zhou, Xingchen, and Rob Jenkins. "Dunning–Kruger effects in face perception." Cognition 203 (2020): 104345.

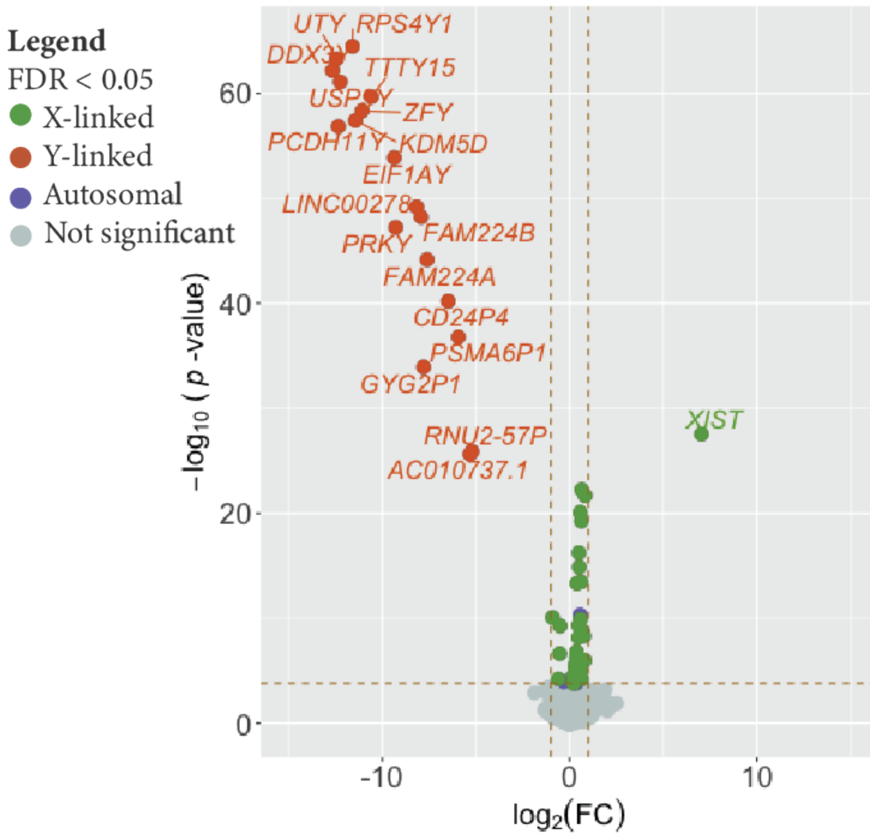
APPENDIX A

FALL 2022-B: ASSESSMENT QUESTIONS

#	Aim	Question, Answers (bold), and Distractors
Q1	Biology	Which of the following is/are true about the placenta? (Choose all statements that are true) **A. The placenta functions as the fetus' excretion, endocrine, and immune systems** **B. The placenta attaches to the fetal body through the umbilical cord and to the maternal body through the uterus** **C. The placenta contains both fetal and maternal blood vessels** **D. The placenta develops from the same fertilized egg that created the fetus.** E. The placenta originates from the maternal body, starting in the uterus, then connects to the fetus through the umbilical cord
Q2	Biology	Which of the following is/are true about the placenta? (Choose all statements that are true) **A. An embryo with two X chromosomes will develop with a placenta with two X chromosomes** **B. An embryo with an X and a Y chromosome will develop a placenta with an X and a Y chromosome** C. All placentas have two X chromosomes because the genetic mother has two X chromosomes (46, XX) D. Although there are sex differences in fetal growth, there are no known sex differences in the placenta
Q3	Biology	Which of the following statements about the sex chromosomes (X and Y) in humans is/are true? (Check all that apply) **A. X and Y chromosomes share evolutionary history and were once homologous autosomes.** **B. The Y chromosome is much shorter than the X chromosome** **C. There are regions of the sex chromosomes that are 98 - 100% identical.** **D. Common sex chromosome variations include Klinefelter syndrome (47, XXY) and Turner syndrome (45, X).** E. **XY individuals can lose their Y chromosome within a proportion of their cells over time as they age.**
Q4	Biology	Which are true about RNA and RNA sequencing experiments? (Check all that apply) A. RNA is more stable that DNA **B. RNA is reverse transcribed into cDNA prior to sequencing** **C. In sequencing experiments the molecules (RNA or DNA) from the cell will be cut into shorter "reads", usually between 75 and 300 nucleotides** **D. The sequencing reads of a gene in an RNA-seq experiment represent a relative quantification of RNA (e.g., relative to other genes sequenced in the sample** E. The sequence reads of a gene in a RNA-seq experiment are a measure of the total number of RNA reads in the cells that were sequenced
Q5	Biology	Choose the best order for processing RNA-seq data to analyze differences in gene expression A. Step 1: [Choose of the following: **Check for initial quality (raw),** Trim low-quality reads, Check for trimmed quality, Align to a reference genome, Quantify reads and generate counts, Perform differential expression analysis]

		B. Step 2: [Choose of the following: Check for initial quality (raw), **Trim low-quality reads**, Check for trimmed quality, Align to a reference genome, Quantify reads and generate counts, Perform differential expression analysis] C. Step 3: [Choose of the following: Check for initial quality (raw), Trim low-quality reads, **Check for trimmed quality**, Align to a reference genome, Quantify reads and generate counts, Perform differential expression analysis] D. Step 4: [Choose of the following: Check for initial quality (raw), Trim low-quality reads, Check for trimmed quality, **Align to a reference genome**, Quantify reads and generate counts, Perform differential expression analysis] E. Step 5: [Choose of the following: Check for initial quality (raw), Trim low-quality reads, Check for trimmed quality, Align to a reference genome, **Quantify reads and generate counts**, Perform differential expression analysis] F. Step 6: [Choose of the following: Check for initial quality (raw), Trim low-quality reads, Check for trimmed quality, Align to a reference genome, Quantify reads and generate counts, **Perform differential expression analysis**]
Q6	Biology	Which of the following are reasons to apply trimming software to reads? (Choose all statements that are true) A. To decrease the total of uniquely mapped reads **B. To remove adapter sequences used in sequencing library preparation.** **C. To remove low-quality reads** **D. To filter out trimmed reads that are less than a specified length**
Q7	Biology	In an experiment you have two conditions, a control condition, and a treatment condition. You have collected multiple samples from the control condition and multiple samples from the treatment condition and sequenced the RNA (RNA-seq). You then apply a normalization step to your RNA-seq data. What is the purpose of the normalization step? A. To ensure the expression distributions of only the control samples are similar across the entire data set B. To ensure the expression distributions of only the treated samples are similar across the entire data set **C. To ensure the expression distributions of all samples are similar across the entire data set** D. To ensure the expression distributions of none of the samples are similar across the entire data set
Q8	Coding	Which are true about R? (Choose all true statements) **A. There are a lot of published statistics and algorithms for biological applications written in R** B. Code written in R is easy to read so you do not need to spend much time adding descriptive comments to your scripts **C. Saving R scripts with the data you produced or analyzed allows you to easily reproduce, modify, and share your protocols for future work** D. R is a coding environment, but not a language E. R is a language, but not a coding environment **F. R is a language and an environment for statistical computing and graphing**

Q9	Coding	The benefits of using the R Markdown format to write your code include which of the following? (Check all that apply) A. R Markdown is the only way to write R code **B. You can print code R Markdown as a report with nicely formatted text, code, and plots** **C. R Markdown is integrated into RStudio to make it easier to work with R code** D. R Markdown only allows you to use specific data types in R making your code simpler to read
Q10	Coding	In a RMarkdown file (.rmd), this symbol is used at the beginning and end of chunks to indicate where to separate code A. >>> B. ### **C.** ``` D. %>%
Q11	Coding	 This is a multi-dimensional scaling (MDS) plot built off the top 100 most variably expressed genes between male and female placentas. Here male is defined as placentas from XY offspring assigned male at birth and female is defined as from XX offspring assigned female at birth. Which of the following is/are valid interpretations of this plot? (Check all that apply) A. There are outliers in both the female and male sample groups on the first scaling component (x-axis) **B. The largest dimension of variation in the top 100 genes is explained by whether a sample came from a male or a female.** **C. Female samples are, in general, more similar to other female samples than they are to male samples, when considering the top 100 most variable genes.** D. The second component of variation (y-axis) separates out male from female samples.

<table>
<tr><td></td><td></td><td>E. The second component of variation shows that samples from males and females exhibit variation, but we don’t know what explains it.</td></tr>
<tr><td>Q12</td><td>Coding</td><td>


This is a volcano plot showing sex differences in expression, with the log2(fold change (FC)) of genes plotted on the x-axis and the p-value plotted on the y-axis. Genes to the right of zero on the x-axis show higher expression in females while genes to the left of zero show higher expression in males. Based on the following volcano plot, which of the following statements are true about the gene UTY
A. UTY has a log2(FC) > 60
B. UTY has a smaller p-value than XIST, and therefore, is more statistically significant
C. XIST has a smaller p-value than UTY, and therefore, is more statistically significant
D. UTY has a log2(FC) < -10
E. XIST has a log2(FC) > 0
F. XIST is an X-linked gene
G. UTY is an autosomal gene
</td></tr>
<tr><td>Q13</td><td>Coding</td><td>When conducting an experiment, we design our hypothesis so that we have a null hypothesis (what is the default expectation for our data) and an alternative hypothesis (what a different explanation for the data). In statistical analyses, one way to determine statistical significance is to compute a p-</td></tr>
</table>

		value. If the p-value you computed is below a chosen alpha threshold (typically 0.05), which would you do (select the correct option) A. Accept the alternative hypothesis B. Reject the alternative hypothesis C. Accept the null hypothesis **D. Reject the null hypothesis** E. Conclude that you cannot reject or accept the null hypothesis
Q14	Coding	The code for the project is in a directory whose full path is /data/project/placenta/RNA-seq/code/. You want to copy this code to your home directory in a directory called /home/user/code/ so that you can make edits on it and run the code with slightly different parameters. Which of the following code will work to copy the code to your home directory? A. `cd /data/project/placenta/RNA-seq/ /home/user/code/` B. `cd /data/project/placenta/RNA-seq/* /home/user/code/` C. `mv /data/project/placenta/RNA-seq/* /home/user/code/` D. `cp /data/project/placenta/RNA-seq/ /home/user/code/` **E. `cp /data/project/placenta/RNA-seq/* /home/user/code/`** F. `cd /home/user/code/ /data/project/placenta/RNA-seq/` G. `cp /home/user/code/* /data/project/placenta/RNA-seq/`
Q15	Professional Development	Author contributions mean different things in different fields. However, authorship does imply contribution to the manuscript. Match the authorship title with its typical definition in our field (Biological Sciences). A. Contributed primary analysis, troubleshooting, writing, and finalizing all components of the manuscript [Choose one of the following: Senior/last author, Acknowledgements, Middle Author, **First Author**, Corresponding Author] B. Person to whom communications about the manuscript (questions, comments) should be directed. [Choose one of the following: Senior/last author, Acknowledgements, Middle Author, First Author, **Corresponding Author**] C. Principal investigator of the lab, conceived the original idea for the research and/or supervised the project, obtained funding and resources. [Choose one of the following: **Senior/last author**, Acknowledgements, Middle Author, First Author, Corresponding Author] D. Contributed specific aspects of the work, writing, supervision, or funding of the project. [Choose one of the following: Senior/last author, Acknowledgements, **Middle Author**, First Author, Corresponding Author] E. People who provided comments, advice, critiques, feedback on the project and/or manuscript. [Choose one of the following: Senior/last author, **Acknowledgements**, Middle Author, First Author, Corresponding Author]
Q16	Professional Development	Which of the following are questions you should ask when reading a scientific paper? **A. What do the author(s) want to know (motivation)?**

		B. What did they do (approach/methods)? **C. Why was it done that way (context within the field)?** **D. What do the results show (figures and data tables)?** **E. How did the author(s) interpret the results (interpretation/discussion)?** F. **What should be done next?**
Q17	Professional Development	Please match the correct description to the term. Subject and what aspect of the subject was studied. Summary of paper: (1) The main reason for the study, the primary results, the main conclusions (2) Why the study was undertaken (3) How the study was undertaken, (4) What was found (5) Why these results could be significant and what the reasons might be for the patterns found or not found. A. Subject and what aspect of the subject was studied. [Choose one of the following: Discussion, Results, Introduction, Abstract, **Title**, Methods and Materials] B. Summary of paper: The main reason for the study, the primary results, the main conclusions [Choose one of the following: Discussion, Results, Introduction, **Abstract**, Title, Methods and Materials] C. Why the study was undertaken [Choose one of the following: Discussion, Results, **Introduction**, Abstract, Title, Methods and Materials] **D.** How the study was undertaken [Choose one of the following: Discussion, Results, Introduction, Abstract, Title, **Methods and Materials]** E. What was found [Choose one of the following: Discussion, **Results**, Introduction, Abstract, Title, Methods and Materials] F. Why these results could be significant and what the reasons might be for the patterns found or not found. [Choose one of the following: **Discussion**, Results, Introduction, Abstract, Title, Methods and Materials]
Q18	Professional Development	Which are options when you are coding or doing research and trying to figure out an answer to your problem? **A. For many programming tasks, others have posted concise, tested code to accomplish those tasks and posted them online** **B. There are multiple websites where you can search for answers, including developer forums like Stack Exchange and Stack Overflow, and even YouTube** **C. To compose a search for how to solve a coding problem, we should include the programming language and the words to indicate the code.** D. To compose a search for how to solve a coding problem one should put the search term in quotation marks **E. To compose a search for how to solve an error in our code, one should put the search term in quotation marks.**
Q19	Personal Feelings	How would you describe your comfort level with using a command line interface to interact with a Linux/Unix command-line style environment? A. Very uncomfortable B. Uncomfortable C. Neutral D. Comfortable E. Very comfortable

Q20	Personal Feelings	How would you describe your comfort level with programming in R? A. Very uncomfortable B. Uncomfortable C. Neutral D. Comfortable E. Very comfortable
Q21	Personal Feelings	How would you describe your level of coding expertise using any programming language?
Q22	Personal Feelings	How comfortable are you asking your peers coding questions in an open class forum? A. Very uncomfortable B. Uncomfortable C. Neutral D. Comfortable E. Very comfortable
Q23	Personal Feelings	How comfortable are you reading and interpreting a scientific paper? A. Very uncomfortable B. Uncomfortable C. Neutral D. Comfortable E. Very comfortable
Q24	Personal Feelings	How comfortable are you in writing a scientific paper? A. Very uncomfortable B. Uncomfortable C. Neutral D. Comfortable E. Very comfortable
Q25	Personal Feelings	What is your experience level doing computational research up until this point? A. I've had little to no research experience B. I have completed a course-based research experience (CURE) C. I currently do computational research D. I have done non-computational research E. I have co-authored published research F. I don't want to do research