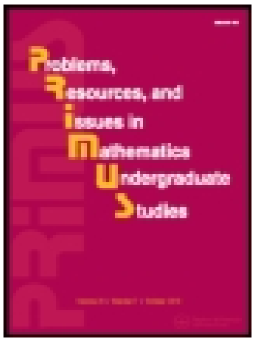

PRIMUS

Problems, Resources, and Issues in Mathematics Undergraduate Studies

ISSN: 1051-1970 (Print) 1935-4053 (Online) Journal homepage: http://www.tandfonline.com/loi/upri20

Mentoring Student Participation in Undergraduate Research: A Case Study of Twists and Turns from Two Perspectives

Ismael Perez, Janet Bowers & Peter Salamon

PRIMUS, 00(00): 1–11, 2016

ISSN: 1051-1970 print / 1935-4053 online
DOI: 10.1080/10511970.2016.1240728

Mentoring Student Participation in Undergraduate Research: A Case Study of Twists and Turns from Two Perspectives

Ismael Perez, Janet Bowers, and Peter Salamon

Abstract: Undergraduate research experiences provide excellent examples of high-impact practices. They rely on inquiry-based learning to provide important capstone experiences for the students. However, they are time-intensive for mentor faculty. In an attempt to scale up our faculty's ability to offer such experiences, we combined a number of projects into one class: the Biomathematics Workshop. This article chronicles the experiences of one student from that class interspersed with comments from his mentor. Our goal is to illustrate some of the twists and turns in his journey and the unexpected but valuable lessons the student (and the mentoring team) learned.

Keywords: Mentoring, undergraduate research experiences, high-impact practices

1. BACKGROUND

Over the past 15 years, the Association of American Colleges and Universities (AACU) has published numerous reports highlighting the need to improve undergraduate STEM programs by including "high-impact practices" both in a student's first few years and as a capstone activity. According to research conducted by Kuh [4] student–faculty research collaboration is one of the highest impact activities that can affect undergraduates' learning and retention in school. In a 2011 report describing these experiences, the AACU authors conclude that "The highest-quality first-year experiences place a strong emphasis on critical inquiry, frequent writing, information literacy,

Address correspondence to Peter Salamon, Department of Mathematics and Statistics, San Diego State University, San Diego, CA 92182-7720, USA. E-mail: psalamon@mail.sdsu.edu
Color versions of one or more of the figures in this article can be found online at www.tandfonline.com/upri.

collaborative learning, and other skills that develop students' intellectual and practical competencies." [1]

Although this call appeals to common sense, the reality is that, for many mathematicians working at colleges and universities, the idea of providing inquiry-based experiences can seem daunting, as the complexity of their research is beyond the reach of most undergraduates - especially those in their first two years of study. This report details how a collaboration process between math and science faculty was designed to overcome time and complexity hurdles by leveraging the work of graduate students (most of whom were pursuing their master's degrees in biology) who were exploring contemporary research problems in biology as part of their thesis work. The first part of this report describes the organization of the Biomathematics Workshop. The second part describes the impact of this course and the continued research efforts during the following semesters from one student participant's perspective, with comments interspersed from his mentor, the third author of this paper.

1.1. Background of the Course

The Biomathematics Workshop was funded through a 3-year NSF grant (which was extended to a fourth and a fifth year) that was designed to integrate mathematical modeling and biological studies at San Diego State University (SDSU). The course ran for four consecutive fall semesters starting in 2010. During the 4 years the course was taken by a total of 32 undergraduate students and seven graduate students. Three of the graduate students were Ph.D. candidates enrolled in either biology or computational science programs. The remaining four were seeking their master's degrees in biology (2), mathematics (1), and mechanical engineering (1). It is important to note that we had a limited population of Ph.D. students from which to choose; our university does not grant Ph.D. degrees, but we do have joint doctoral programs with several local universities.

Potential participants were recruited by email and word of mouth to undergraduate students, graduate students and faculty. The emails to biology graduate students and faculty solicited problems they were working on that could benefit from quantitative models. The criteria used to determine which projects would be included were: (i) the appropriate level of mathematical complexity; (ii) support from the research group suggesting the program (i.e., the availability of project personnel); and (iii) the project's potential to interest workshop students.

One of the most gratifying aspects of the program is that of the 32 undergraduate students who enrolled in the course during its 4-year run, 26 of them continued to work on their respective projects after they had finished the course. Most received compensation through other programs on campus such as the Minority Access to Research Careers (MARC) program, Minority Biomedical Research Support, Scholarships in Science Technology, Engineering and Mathematics, and the McNair Scholars program. Some were pre-med majors and wanted the experience of doing research, writing papers, and presenting at conferences. In addition to funding some of the participants, the grant was used to support a buy-out for the two faculty members (Salamon, who is an author of this paper, and one other from biology).

As the course was designed to provide an early learning experience, the only prerequisite was completion of two semesters of calculus and a GPA of 3.0 or higher. After an iteration of the course, both faculty members realized that using undergraduates to model real-world problems involved an overwhelming amount of mentoring time. In addition, the fact that students from previous years continued their work meant that these two faculty members were mentoring more and more students each year. Therefore, in ensuing iterations of the course, several additional graduate students were recruited.

Ismael, whose account of the experience is presented below, joined the workshop course during its second year, fall of 2011, when he was a sophomore. This revised workshop began with the graduate students (and undergraduates from the previous year) describing their projects and the mathematical models that were needed. Next, undergraduates chose the topic they wanted to work on, and subgroups containing teams of undergraduate and graduate students were formed for each of the chosen topics. Subgroup meetings were used to discuss progress and methods that pertained to the individual topics, and the tools needed to model the situations. Mandatory weekly meetings were also scheduled with all students and faculty to hear reports on progress, problems, and issues.

After Ismael completed the course in the fall of 2011, he continued his research throughout his junior year (2012–13) and presented his findings at several conferences. The goal of this article is to describe how this opportunity affected his own personal growth, and to offer his story as a case study of the unexpected, but valuable lessons that engagement in high-impact practices such as this provides undergraduates. In Ismael's case, his research problem took many unexpected twists and turns, but ended up developing a new application of perturbation theory to parameter identification. In what follows, we present his reflection on the experience interspersed with comments from his mentor.

1.2. Ismael's Background (in his own words)

As a first generation college student, a research career was far away from my scope of possible professions. The summer before I started my first semester in college, I participated in a minority bridge program. The program introduced participants to possible careers in science and engineering fields. This led me to the research participation that started in the Biomathematics Workshop class. In this course, I was introduced to several research projects that students had worked on in the past and was invited to select one of the projects. The students on a project worked with each other and practiced presenting their projects at weekly lab meetings. As a result, the course provided an introduction to research, and also served to improve our public speaking skills. Oral communication was difficult because of stage fright, which the practice sessions with the group helped overcome. Written communication was also needed for creating posters, Beamer presentations, and manuscripts. Participating in undergraduate research provided the opportunity to develop these important communication skills.

I did not know what to expect from doing research in mathematics. My biggest surprise was that I was learning about things that I had not learned from any of my science and math courses. Teaching myself the necessary material for my research project was the most beneficial, and at the same time, the hardest part of doing research.

2. THE PROBLEM

Ismael's chosen problem concerned finding values of the constant parameters k_i in the linear system of differential equations (1)-(4) in Figure 1, which matched experimental values of A, B, C, and E. Given values of the k_i variables, the equations can be integrated (in closed form) and the sum squared error between the solution curve and measured data needed minimization as a function of the k_i. It thus appeared to be a problem within reach of the student population.

In presenting this problem, the graduate student who originated the problem and an undergraduate from the previous year who had worked with him (both of whom were biology majors) explained that they had so far managed to find the values of only two out of the five parameters.

Ismael:

My research project was applying perturbation theory to a system of linear differential equations. Although I did not realize this at the time, doing perturbation on a linear system is something that at first glance sounds crazy to a professional mathematician or physicist. The usual reason to use perturbation

$$\frac{dA}{dt} = -k_1A + k_{-1}B \quad (1)$$

$$\frac{dB}{dt} = k_1A - (k_{-1} + k_2)B + k_3C \quad (2)$$

$$\frac{dC}{dt} = k_2B - (k_{-2} + k_3)C \quad (3)$$

$$\frac{dE}{dt} = k_3C \quad (4)$$

Figure 1. The chemical reaction. $A \leftrightarrow B \leftrightarrow C \rightarrow E$.

theory is to make a problem linear. To use this theory on a problem that is linear to begin with seems futile.

Mentor comment:

Ismael had not had a differential equations class at this point. Although our Engineering Calculus sequence does cover the process of solving separable differential equations, being thrown into a situation that demanded working with a system of them was certainly daunting. He was guided through the necessary learning at this roadblock by the math faculty member, through discussions in the subgroup meetings and through guided readings on his own.

With regard to the approach, the sad fact was that Ismael's statements above were perfectly accurate – doing perturbation theory on a linear problem was looked at with skepticism by all who viewed it, except the faculty member (Salamon) who suggested it. This made it very difficult for Ismael to maintain credibility with his audience at presentations of his work. Remember though, that our goal was to identify more coefficient values. It turns out that doing perturbation theory on a linear system does help identify more parameters, but the skepticism that Ismael met was a serious problem for his developing self-esteem. This "PR issue" remained a thorn until the end of the project, and is a likely feature of many genuine research efforts.

Ismael:

The system of linear differential equations I was working with described the chemical kinetics of site-specific recombination by the bacteriophage lambda. This is a mechanism that viruses such as bacteriophage lambda use to insert their DNA into the DNA of their *E. coli* host [5]. Our task was to fit a given system of differential equations to data and thereby find values of the parameters of biological interest, i.e., I needed to identify the constants in a linear system of differential equations. This task itself was abstract, since I had never tried to use any computer software involving mathematics. Therefore, I soon learned to use Matlab and Maple.

Mentor comment:

This was one place where the other undergraduate students who had taken the *Calculus for Biology* course could really help him. Although the standard engineering calculus course that Ismael had gone through does not cover these topics, our Life Science Calculus sequence does cover differential equations and includes many computer lab activities that fit unknown coefficients for them [6]. Help from his peers made it possible for Ismael to catch up with the previous year's results quickly. It also helped him acquire the computer skills, although much of this had to be done on his own.

Previous attempts by the research group to identify the coefficients in the kinetic equations had run into difficulties [3, 8]. As I learned later, this was due to limited observability – two of the species [B and C] in the reaction were not measured separately; the experimental data included only their sum. In addition, the reaction that interconverts these two species approaches equilibrium on a faster time-scale than the time-scale of our observations. If this fast reaction is always truly in equilibrium, we cannot see the rate constant for this equilibration. However, a weak signal from the "fast" reaction can be seen in the first few seconds. To focus in on this signal, we studied the perturbed system expanded in time from very particular initial states hoping to obtain a better identification of the rate constants.

Ismael:

When I started the research project, I had never heard of perturbation theory, but applying what I had learned about Taylor series expansions in my Calculus courses, I was able to slowly understand the concept involved. The problem then became solving for the zeroth- and first-order terms in the perturbation expansion. I next ran into an unanticipated setback. I had the zeroth-order solution, but was not able to use it because the solution required the values for the initial concentrations of the indistinguishable chemical species. As a result, I was not able to use the zeroth-order solution that I had found, and was forced to continue on to obtain the first-order terms for the perturbation expansion of each state variable.

In finding these terms of the expansion, I faced a more challenging calculation than my courses had prepared me for. Knowing only paper and pencil methods, I kept stacks of papers of my calculations, and continued to go over them to make sure I did not make any mistakes. Unfortunately, the first few times through, I did find mistakes and got frustrated, because I had to go back and do the calculations again and again. This was a difficult part of the research project, not just because it was a lengthy calculation, but also because it was my first time doing mathematics for which no one knew the right answer, and which needed to be checked by methods I had to think up. I had to come up with many checks along the way, but the final check was to plot my solution and compare it to the unperturbed solution of the system of linear differential equations.

After successfully obtaining the perturbation solution, I fitted the model to the data and faced a new setback. Sadly, even the perturbation approach was not able to give us stable parameter identification. Later I was able to show that this was due to the low precision of the data available from the experiments. I once again felt discouraged about the research project. I had expected our perturbation solution to work, but eventually came to understand that such setbacks are not atypical and are very much a part of the research process.

Mentor comment:

This was indeed a serious setback and did not help morale for the project. It also coincided with the end of the formal course. I was not ready to give up, however. Suspecting that the noise was responsible for our failure, I suggested that Ismael try taking a known solution generated with some values of the parameters and see if his computer code could reproduce those same parameters, i.e., try synthetic data.

Ismael:

Next I turned to synthetic data to demonstrate that our method could indeed help identify parameters, at least for noise-free data. I therefore examined the Hessian of the sum-squared error as a function of the parameters for both the full and the perturbed models. To obtain the components in the Hessian matrix for the sum-squared error, I utilized a numerical approximation to the second-order partial derivatives. This whole approach was foreign to me, because I had no previous experience in numerical techniques.

Mentor comment:

Ismael had taken our lower-division Linear Algebra course and the numerical linear algebra he needed was relatively easy to learn. On the other hand, the probability theory needed to add synthetic noise was more problematic. These skills were acquired over the next 2 years of weekly meetings and one-on-one sessions along with guided readings.

Ismael:

The Hessian matrix gave me the eigenvalues and eigenvectors of both the full model and the first-order perturbation model. The eigenvalues of the Hessian matrix describe the curvature at the minimum. Our results showed that the smallest eigenvalue of the first-order perturbation model was an order of magnitude greater than the smallest eigenvalue of the full solution, which meant that the model did a better job at identifying the desired parameters. This, at long last, was a real success and vindicated the use of perturbation techniques for this problem.

The last part of my project was to add noise to the synthetic data and compute the resulting distributions for the smallest eigenvalues. Once again, I did not have any background in the task that I needed to complete. My courses had not included probability theory, but I learned that you have to go where the problem leads you. I was successful in again showing that the first-order perturbation expansion did a better job at identifying the minimum for two of our parameters. I was also able to determine the maximum noise level for which this was true. This noise level was about 2%, well below the 5% noise level in the real data. On the other hand, it promises to be within reach of the newly designed microfluidic realization of the experiment inspired by my findings.

My active participation in the undergraduate biomathematics research group allowed me to present my research first at regional MAA meetings and later at many conferences including the Microbiology Group Annual Meeting, the SACNAS National Conference, the Joint Math Meetings; the Pacific Coast Conference, and three consecutive student research symposiums at SDSU. My active involvement in the Interdisciplinary Training for Undergraduates in Biological and Mathematical Sciences (UBM) research group also made me a strong candidate when I was applying to the MARC honors program. This program provided academic and financial support for my last 2 years at SDSU and helped prepare me for doctoral programs.

3. CLOSING REMARKS

Definitions of what constitutes high-impact practices and inquiry-based learning are necessarily broad. For example, conducting inquiry-based learning requires more than just having a teacher or student ask and answer a series of scaffolded questions. Instead, this type of learning occurs "…when individuals attempt to convert information and data into useful knowledge." [2] Was Ismael pursuing inquiry-based learning throughout the project? The goal was very clear: find the values of the five constants k_i in equations (1)-(4). However, the path to find these parameter values was definitely not clear. As shown in Figure 2, Ismael's inquiry process led to problems needing techniques that forced him to learn a good deal of mathematics that he had either not yet encountered or that he had learned, but had not recognized as applicable to real-world applications. It is clear that Ismael's work developing questions, working with his team to create computer programs, and using these explorations to cement his understanding of the problem and his solution all meet the definition of a high-impact inquiry-based learning project.

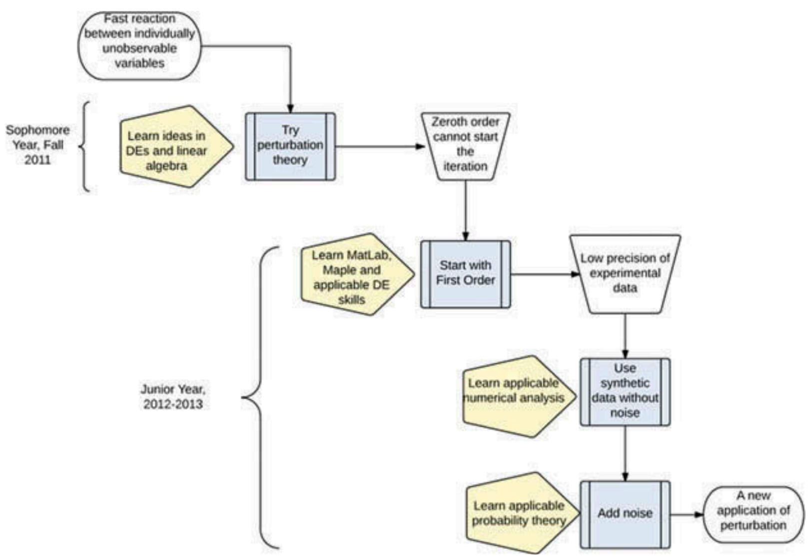

Figure 2. The twists and turns in Ismael's inquiry process.

Should we have required more prerequisites? Maybe. This is the old, familiar quandary: if students start too early, they are unprepared but have time to do a serious project. However, if they start later on in their college careers, they are better prepared, but must finish within a very tight schedule. This report illustrates that one possible way to address this conundrum is to engage more experienced students to share their expertise.

Ismael's reflections provide insights that help illuminate the types of "impacts" that high-impact practices can have. These include the following.

3.1. Real-World Data is "Messy"

Working with actual data differs from working with contrived problems found in many textbooks. Business leaders have stated that when they hire new graduates, they look for candidates who understand that models do not always fit data perfectly, especially the first time around.

3.2. Persistence is Critical

Solutions to real-world problems are not known, and perseverance and personal dedication are required. As Ismael said, "[This] was my first time doing mathematics for which no one knew the right answer and which needed to be checked by methods I had to think up." This insightful comment reflects an irony of current teaching. Although young children ask all types of questions, older students who have become accustomed to the school routine cease asking questions on their own. Moreover, they come to believe that answers should be found very quickly, and persistence is neither necessary nor rewarded. Experiences such as this can challenge the prevailing didactic contract [2] and empower students to persist in their research.

3.3. College Courses only Provide Foundations

One of the most remarkable aspects of this case study is that Ismael took the initiative to learn the math content and programming skills needed to solve the problem. Moreover, he realized that content courses can never cover all of the material that one will need in the future. The goal of an undergraduate education should be to learn *how* to learn and *how* to work with others to find answers. Although mathematical skills are critical to today's STEM workforce, many math majors do not work closely with students in other disciplines. In fact, "The educational offerings of typical departments in the mathematical sciences have not kept pace with the changes in how the mathematical sciences are used." [7, p. 124]. This case study provides one example of how the departments of biology and mathematics and statistics can leverage real-world research to provide more enriching material for the math majors, along with more mathematically complex modeling for the science majors.

3.4. Personal Confidence is Key

Feelings of discouragement are a natural part of any inquiry process. Science is a process, not a quick journey, and those who look at “roadblocks” as road signs rather than stumbling blocks are more likely to succeed. As Ismael’s mentor explains, part of the difficulty of the task was persevering when other “experts” discouraged the approach. Ismael’s reflection communicates a growing personal confidence in his ability to communicate with his team and his mentor.

In summary, these comments (and the reflections from Salamon) indicate that the value of this experience extended beyond the often-measured outcomes of persistence in a STEM major or improvement in final grades in a course. In fact, they indicate that Ismael experienced the type of real-world problem-solving and communication skills that employers consistently say are lacking in most recent graduates.

Although the value of such experiences for undergraduates is clear, faculty hesitate to get involved with such programs, fearing their research area to be too technical for undergraduates. As Ismael’s example shows, the undergraduate research participation need not be in the faculty member’s area. A math clinic built around research problems brought in from across campus can be very accessible. Ismael’s problem, fitting data to a known system of linear differential equations appeared, at the outset, to be very doable. Nonetheless it took several twists and turns, as research problems often do. Such surprises are probably unavoidable in such scenarios and provide the most valuable insights of all: Personal confidence and persistence are the true keys to success.

4. POSTSCRIPT

As a result of Ismael’s experiences, he was able to land a valuable internship at a National Lab and is now completing his MS degree in applied mathematics. He has a job offer, and is currently applying to several Ph.D. programs.

FUNDING

We gratefully acknowledge support from NSF UBM-0827278, Undergraduate Biomathematics Program and SDSU MARC Program NIH/NIGMS Grant #5T34GM008303-23.

REFERENCES

1. Association of American Colleges and Universities. 2011. *The LEAP Vision for Learning: Outcomes, Issues, Impacts, and Employers' Views*. Washington, DC: Association of American Colleges and Universities.
2. Educational Broadcasting Corporation. 2004. Concept to classroom. Workshop: Inquiry-based learning. http://www.thirteen.org/edonline/concept2class/inquiry/. Accessed 23 February 2015.
3. Esquivel, S. 2013. Kinetic studies of phage lambda integrase mediated Holliday Junction resolution. Undergraduate honors thesis, San Diego, CA: Department of Biology, San Diego State University.
4. Kuh, G. D. 2008. *High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter*. Washington, DC: AACU.
5. Lodish, H., A. Berk, M. Krieger, M. P. Scott, A. Bretscher, H. Ploegh, and P. Matsudaira. 2007. *Molecular Cell Biology*. New York: Freeman.
6. Mahaffy, J. and A. Chavez-Ross. 2009. *Calculus: A Modeling Approach for the Life Sciences (Volumes 1 and 2)*. San Diego, CA: Pearson Custom Publishing.
7. National Research Council. 2013. *The Mathematical Sciences in 2025*. Washington, DC: The National Academies Press.
8. Yi-An, L. 2013. Interactions of proteins and small molecules with Holliday Junction. MS thesis, San Diego, CA: Department of Biology, San Diego State University.

BIOGRAPHICAL SKETCHES

Ismael Perez earned his BA degree from San Diego State University in May, 2014. He graduated cum laude with a major in mathematics with an emphasis in computational science.

Janet Bowers is an associate professor at San Diego State University. She is a mathematics educator who has an interest in improving the instructional experiences of undergraduate students. She received her Ph.D. from Vanderbilt University

Peter Salamon is a mathematics professor at San Diego State University. He has published well over 140 mathematical articles related to biomathematics, thermodynamics in finite time / geometrical thermodynamics, and optimization and mathematical modeling. He received his Ph.D. from University of Chicago.