CHOOSING A RESEARCH QUESTION IN APPLIED MATHEMATICS, FROM MENTORS TO NOVICES

Carlos W. Castillo-Garsow
Eastern Washington University
ccastillogarsow@ewu.edu

Based on observations made during the project proposal week of an undergraduate research program in applied mathematics, this paper explores the role of faculty mentors in guiding students to developing a research question and accompanying a model. Results suggest that students should be pushed to become experts in the background subject matter, while mentors take the lead mathematically. Key skills for developing a research question include: defining the temporal/spatial focus, exploring broader impacts of the work, and anticipating possible mathematical results to help define the question. Constant dialogue on both sides about the scientific mechanisms informing mathematical choices was critical to model development.

Mathematics as it is taught in secondary and post-secondary classes differs greatly from mathematics as it is practiced by professionals and from the needs of partner disciplines (Ganter & Barker, 2004; Ganter & Haver, 2011; Lewis & Powell, 2017). Reports from industry and professional societies repeatedly emphasize the importance of professional skills in communication, collaboration, problem solving, mathematical modeling, and creativity on top of a solid foundation of procedural skills and coherent mathematical understandings (Bliss et. al., 2016; Ganter & Barker, 2004; Ganter & Haver, 2011). These soft-skills do not come free with fluency in mathematical skills or sophisticated mathematical understandings.

The skill of interest to this report is described by Bliss et. al. (2016) as "Distilling a large ill-defined problem into a tractable question," (p. 72) which I will call "developing a research question." Smith et al. (1997) found that graduate students in mathematical biology struggle with developing research questions that are both biologically interesting and mathematically tractable.

Mathematical modeling education research has little to say on the development of this skill. Research in this area typically focuses on students working pre-chosen tasks (Bliss et. al., 2006; Gravemijer, 1994; Lesh & Doerr, 2003). Sometimes these tasks are quite open, and students go through cycles of model development; however, in assigning a task to students, there are constraints placed on students as part of intentionally guiding the students' conceptual development (ibid). It has been argued in the past that these constraints limit students' experiences in developing their own research questions (Castillo-Garsow, 2014; Castillo-Garsow & Castillo-Chavez, 2015).

Camacho et al. (2003) found that choosing one's own project and research question creates situations in which students take the lead in researching topics far outside a mentors' area of expertise, essentially reducing the mentor to a role of consultant rather than leader. These unique situations create opportunities to identify key components of soft-skills such as developing a research question. For example, in studying student-chosen projects, Smith et al. (1997) found that students' and mentors' goals for modeling, as well as their beliefs about the value and purposes of modeling, impacted their ability to develop a research question that resulted in a scientifically relevant and tractable model.

This study follows a single group of undergraduate students in the process of developing a topic of their own interest into a research question and accompanying mathematical model. Because the students in this project worked in close and constant collaboration with both graduate and undergraduate mentors, we can see how mathematicians at different stages of their career view the task of developing a research question and accompanying model.

The purpose of this study is two-fold: (a) to begin the process of identifying specific goals, skills, and values that are critical to developing a research question and model in applied mathematics, and (b) to provide guidance to mentors of student-led applied mathematics projects by identifying effective, transferable interventions.

Methods

This study occurred in the fifth week of an eight-week summer REU in mathematical biology. Prior to this study, the students had taken a three-and-a-half-week course consisting of lecture, computer lab work, and textbook exercises in dynamical systems. Following this course work, students self-recruited into groups of three to five, and chose a topic of interest. During the fifth week, students made daily presentations on their topic to a panel of faculty and graduate mentors who provided feedback. In the final three weeks of the program, students completed the analysis of their model and wrote a technical report on their project. Four groups of students chose to participate in the study, and these preliminary results are from the analysis of the first group. This group was chosen to analyze first because the project was judged by participating mentors to be the closest to a typical project in the REU, and because the success of the project could be determined by publication in a prominent journal. The citation for the publication is omitted to protect the privacy of the participants.

The group of students in this study was formed of five undergraduate students who chose to construct a model for controlling a disease that is transmitted between multiple species of animals. They made six presentations during proposal week to a panel of faculty and graduate mentors. One presentation (five) was cut short by the mentors who did not believe that their feedback was necessary at that time. Thirteen of these mentors made comments on the students' work. Each proposal conference was video and audio recorded, and the audio recordings were transcribed. These transcripts, along with the technical report written at the end of the summer program were taken as data for analysis. The published version of the students' project, which resulted from unrecorded collaboration after the program ended, was omitted from this study.

Transcripts were initially open-coded and then axially coded (Strauss & Corbin, 1990) using qualitative data analysis software. This coding provided a visualization for how the content of the presentations and the priorities of the participants changed as the project developed. However, this initial coding did not give a sense of the impact that mentors' comments had in influencing the direction of the project, so a second analysis was performed. In this second analysis, mentor comments were isolated from the transcript, and each mentor comment was coded individually for its content, and then in the context of the transcript and the final paper with an eye to how the content of that comment was revisited in students' future work. Over the six daily presentations, 269 comments from 13 mentors were isolated, coded, and analyzed.

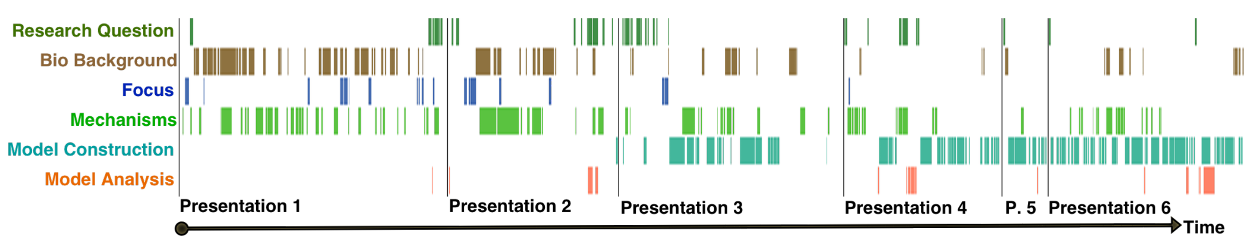
Results

Initial Coding

The initial coding provided a sense of the structure of the presentations. Figure 1 show the top level codes for the topics being discussed by both students and mentors during the presentation. Topics generally fell into nine broad categories in this analysis, only six of which are discussed in this paper due to space limitations. The first code was related to discussions of the students' research question, such as asking students to present their research question or discussing what makes a good research question (ex: "I'm curious what a good measure is for whether [your intervention is] effective"). Background codes referred to discussion of the biological background situation in general, but not specific to developing the model, such as the initial literature review into the behavior of the disease or the life cycle of an animal (ex: "But if you put those same surface proteins on the bacteria that they do produce an immune reaction to, then they'll build up antibodies with that initial infection, they will start fighting off the bacteria"). Focus codes referred to codes that were about defining the problem space to a specific geographic, demographic, spatial, or temporal region (ex: "the other data sets were from Indiana, which is a very different ecological setup than in this area"). Mechanism codes referred to key processes that informed model development, such as identifying stages that individuals pass through, choosing variables, describing the precise way that individuals counted by those variables interacted, and proposing specific intervention strategies (ex: "Yes, so every time they are moving onto the next stage… that's when they can pick up the bacteria"). Model codes referred to the development of the equations and corresponding flow diagrams to be used (ex: "it should be NI + NS. Infected and susceptible"). Lastly, model analysis codes referred to mentors anticipating the results that students might get (ex: "Regardless of whether these treatments are able to reduce R0 below one, they will reduce the endemic prevalence somewhat").

Figure 1

Presence or absence of a code in the transcript over time



Note: timeline of topics discussed by both mentors and students during the students' six presentations. Color indicates the topic is being discussed at that time, while white indicates the topic is not being discussed at that time. Black vertical lines separate individual presentations.

Three topics were prominently discussed in every session (Figure 1): the research question (dark green), the background biology (brown), and the specific mechanisms that would inform model construction (light green). Two days of presentations were devoted to these topics before the research question is defined and model construction began, and during model construction, these topics – particularly mechanisms – continued to be revisited.

Much of the early discussion in defining a research question focused around defining the scope of the project (Figure 1, dark blue). Prior to presentation three, a prominent feature of the discussion was placing specific bounds on the scale of the study: the geographic location to be modeled, the specific populations to be studied, and time scale to be used. These questions provided direct guidance in refining the research question.

An unanticipated result is that the discussion of model analysis (Figure 1, bright orange) preceded the development of the model (cyan), and even the final determination of the research question itself (dark green). Discussion of possible methods of analysis created a common language that the mentors used to discuss the research question in an unfamiliar topic. By discussing potential mathematical results, the mentors helped guide the research question to being one that could be defined mathematically in the form of a model.

The Second Coding

The second coding isolated mentor comments and related the impact of each mentor's comment on the project to the coded content of the comment. Mentor comments were coded by content, and then the impact of each mentor's comment was coded in three ways (Duration, Fidelity, and Direction). Duration was coded as Final, Local, or None depending on if students responded to comments in the final paper, in presentations, or not at all. Fidelity was coded as Pivotal, Direct, or Tangential, depending on whether students based a key aspect of the project on the mentor comment, followed feedback faithfully, or made changes that were merely related to the feedback. Direction coded the novelty of the mentor's suggestion itself as Identical, Similar, Distinct, or Novel, depending on if the comment was something students had discussed before in presentation, related to an idea students had discussed in presentation, different from something students had discussed in presentation, or a new idea that students had never discussed in presentation. Coding was based only on the evidence available within the presentations and the final paper. It is possible that students may have thought of an idea coded Novel and never presented it, but these possibilities were not a factor in coding.

For examples of codes, the mentor comment "One box can cover how big an area?" was given a code of Similar because students were planning a spatial model, Direct because students followed up and found an answer, and Local because students did not include space as a factor in their final model. The mentor comment "How much are these boxes?" was coded as Novel because students had not proposed to look at cost prior to this question, and Final because costoptimization played a role in the final paper. This comment was coded as Direct rather than Pivotal because students initially only answered the question. The following day the comment "you could impose a cost structure on it" was coded as Pivotal, as students changed their research question after this second comment, but Similar because students had already discussed cost. For an example of Tangential, several mentors suggested that students investigate the feeding behavior of the animal, and students did look for papers on this topic, but never incorporated the idea into their project. Comments that most reliably impacted the final paper were questions that asked students to define the research question or the focus of the project more precisely, and comments that assisted students with model development or choosing a methodology. Comments that received no response from students were ones that requested changes to the direction of the research question or the focus, or asked students to make or explain decisions about model development.

A look at the relationship between Direction and Fidelity showed a similar story. Comments that helped students do what they were already going to do (Identical or Similar Direction) were followed faithfully (Direct Fidelity) 78% and 67% of the time (Identical and Similar respectively). Comments that asked students to make drastic changes (Distinct Direction) were unpredictable, resulting in 43% Pivotal Fidelity, 33% Direct Fidelity, and 25% Tangential Fidelity (after rounding). Comments that introduced new ideas (Novel Direction) were either comments that resulted in radical changes to the project (33% Pivotal Fidelity), or were followed only in a perfunctory way (50% Tangential Fidelity).

Mentor comments that students took a pivotal advice were mentors explaining mechanisms, asking students to explain mechanisms, or mentor suggestions for methodology or analysis. Comments students followed directly were asking students to explain model development decisions, or suggesting specific changes to the model. Students compromised (Tangential Fidelity) on requests to do additional background research or find more data.

Comments that developed ideas students haven't thought of (Novel or Distinct Direction) included suggesting or providing additional background, asking students to explain mechanisms, or suggesting the impact of their research (how it might affect others). Comments that drove students deeper into existing directions (Similar or Identical Direction) included asking students to decide on a model or analysis, or asking students about the impact of their research.

Discussion

In broad strokes, mentors focused on asking questions about the background situation (which they were less knowledgeable of), and provided direct suggestions about the mathematical construction of the model (which was within their area of expertise). The timeline of the presentations also followed a trajectory of background to research question to model. However, a deeper dive into the data identified skills and mentor actions that were not immediately obvious.

The data analysis identified specific ways in which mentors guided students in the development of a research question: (a) mentors pressed students to precisely define the spatial, temporal, and demographic focus of the study, (b) mentors suggested possible impacts or implications for the research, and (c) mentors anticipated possible mathematical analysis or results, which helped guide the choice of research question.

Students appreciated feeling in control of their project. Mentor comments that helped students to do what they already planned to do were followed much more faithfully. Mentor comments that requested large changes to the project were sometimes of pivotal importance, but more frequently were ignored, or partially followed. The difference was the background expertise of students and the mathematical expertise of mentors. Students were reluctant to make decisions about the mathematical construction of the model, but welcomed the opportunity to direct their own project and be experts in the background subject matter.

Lastly, the discussion of mechanisms played a key role in every aspect of research question and model development. Mentors directly pressed students for deeper scientific understanding until that it could be made mathematically precise. This skill of examining how the background process worked, and choosing an expression to describe that process was critical for model development and was exercised consistently throughout all the presentations.

A limitation of the study is that of the 269 comments made by mentors, slightly over half (141 comments) received no response from students. Part of the reason for lack of response of students is the large number of mentors who frequently talked over each other or interrupted, giving the students insufficient time to respond. It is impossible to determine how students might have responded to these comments that were interrupted.

References

Bliss, K., Fowler, K., Galluzzo, B., Garfunkel, S., Giordano, F., Godbold, L., & Zbiek, R. (2016). *Guidelines for assessment & instruction in mathematical modeling education.*

Consortium for Mathematics and Its Applications; Society for Industrial and Applied Mathematics.

Camacho, E. T., Kribs-Zaleta, C. M., & Wirkus, S. (2013). The mathematical and theoretical biology institute-a model of mentorship through research. *Mathematical Biosciences & Engineering*, 10(5&6), 1351.

Castillo-Garsow, C. W. (2014). Mathematical modeling and the nature of problem solving. *Constr. Found*, 9, 373-375.

Castillo-Garsow, C. W. and Castillo-Chavez, C. (2015). Why REUs matter. In Peterson, M. A. and Rubinstein Yanir, A., (Eds.), *Directions for Mathematics Research Experience for Undergraduates*. World Scientific.

Ganter, S. L., & Barker, W. (2004). A collective vision: Voices of the partner disciplines. The Curriculum Foundations Project, 1-13.

Ganter, S. L., & Haver, B. (2011). Responding to the recommendations of the Curriculum Foundations Project. *Partner discipline recommendations for introductory college mathematics and the implications for College Algebra*, 39.

Gravemeijer, K. P. (1994). *Developing Realistic Mathematics Education (Ontwikkelen van realistisch reken/wiskundeonderwijs)*. Utrecht CD β Press.

Lesh, R. E., & Doerr, H. M. (2003). *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching*. Lawrence Erlbaum Associates Publishers.

Lewis, M., & Powell, J. (2017). Yeast for mathematicians: A ferment of discovery and model competition to describe data. *Bulletin of Mathematical Biology*, 79(2), 356-382.

Smith, E., Haarer, S., & Confrey, J. (1997). Seeking diversity in mathematics education: Mathematical modeling in the practice of biologists and mathematicians. *Science & Education*, 6(5), 441-472.

Strauss, A., & Corbin, J. (1990). *Basics of Qualitative Research*. Sage publications.