Professional Development Supports for Teaching Bioinformatics through Mobile Learning
2021 DRK-12 PI Meeting Poster

Jooeun Shim, University of Pennsylvania

Overview

In this an Early Stage Design and Development study where we are co-constructing and field testing multidisciplinary high school units organized around bioinformatics concepts, data literacy, and mobile learning. The overarching goal of the project is to help create an engaged population of informatics-informed students who are capable of critically analyzing information and able to solve local problems related to their health and well-being.

https://youtu.be/AEwLVtvV-CQ
Teacher Interview (1’ 37)
Central Curricula Activities

Air quality data collection sensor and mobile phone app interface

Students collecting indoor and outdoor air quality data

Screenshots of collected air quality data and crowdsourced data visualization tools on project website (k12bioinformatics.org)

Screenshots of sample data visualization tools used for comparison with EPA data
Goals and Research Objectives

Needs
• High quality science teacher professional development and science education
• More integration of bioinformatics units in high school courses
• More inclusive STEM programs for underrepresented populations

Bioinformatics
• An interdisciplinary field that combines aspects of computer science, mathematics, and statistics to collect, store, manage, analyze, and interpret biological data

Data Literacy
• Ability to understand, use, manage, and communicate data for multiple purposes (Gebre, 2018)

Research Objectives (Slide 3)

#1: PD to develop teachers’ instruction in:
   a) core bioinformatics competencies
   b) data literacy
   c) mobile learning

#2: Work with teachers to co-construct a bioinformatics unit for high school students

#3: Collect data on classroom instructional challenges and the influencing factors including teachers’ content knowledge and skills, and population and school level variables

#4: Collect data on student learning and participation outcomes in a) core bioinformatics competencies; b) data literacy; and c) mobile learning

Theoretical Frameworks

Best practices in teacher professional development

1) Focus on **disciplinary content**, both the concepts and pedagogies;

2) Addressing how teachers learn through **active learning** and sense-making;

3) Enabling **collaboration** among teachers;

4) Using **models** of effective instruction;

5) Offering coaching and **expert support**;

6) Dedicated time for **feedback** and **reflection** on practice; and

7) **Sustained duration** of PD participation.

---

Year 1 (2019-2020)

- 3 week **in-person** summer PD workshop
- 3 Biology and 3 Environmental Science Teachers
- School Demographics
  - Science proficiency: Range = 2% - 100%; Mean = 38%
  - Ethnicity (% Minority): Range = 61% - 98%; Mean = 82%
  - Income (Free or reduced lunch): Range = 98% - 100%

<table>
<thead>
<tr>
<th>Week</th>
<th>Details of PD Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Comprised instruction in the various components with a majority of time devoted to working with teachers on bioinformatics, data literacy, and mobile learning content and pedagogical content knowledge.</td>
</tr>
<tr>
<td>2</td>
<td>Required teachers to sample portions of the research team constructed PBL unit; to work together in small teams to make detailed lessons plans with adaptations for their local populations; and to plan how to integrate these activities into their standard curriculum.</td>
</tr>
<tr>
<td>3</td>
<td>Asked teachers to pilot core lessons with a small set of high school students. In addition to requiring teachers to collaborate, we included PD designs that met all the requirements for effective PD as outlined by Darling-Hammond et al. (2017). For example, we dedicated time for discussion and reflection on how project activities would be carried out in practice.</td>
</tr>
</tbody>
</table>
Year 2 (2020-2021)

- 4 week **online** PD workshop (due to COVID)
- 4 Biology and 4 Environmental Science Teachers
- School Demographics
  - Science proficiency: Range = 14% - 96%; Mean = 33%
  - Ethnicity (% Minority): Range = 33% - 97%; Mean = 76%
  - Income (Free or reduced lunch): Range = 98% - 100%

- Design considerations for online PD experiences
  - Utilizing multiple representations of learning content (e.g., pairing audio and video with accompanying text and visualizations)
  - Virtual community building media activities (e.g., whatsapp, monthly synchronous meetups)
  - Social capital building activities (e.g., discussion forums, Google Classroom)
  - Expert peer facilitation (e.g., Cohort 1 teachers facilitating Cohort 2 teachers)
# Problem Based Learning Curriculum

## Scope and Sequence of the Unit

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction to Problem-Based Learning</td>
</tr>
<tr>
<td>2</td>
<td>Introduction to Bioinformatics</td>
</tr>
<tr>
<td>3</td>
<td>Introduction to Genotype and Phenotype</td>
</tr>
<tr>
<td>4</td>
<td>Introduction to the Exposome and Air Pollution</td>
</tr>
<tr>
<td>5</td>
<td>Privacy and the Ethics of Data</td>
</tr>
<tr>
<td>6</td>
<td>Reasoning About Data</td>
</tr>
<tr>
<td>7</td>
<td>Introduction to Data Analysis Tools</td>
</tr>
<tr>
<td>8</td>
<td>Analyzing an Air Quality Data Set</td>
</tr>
<tr>
<td>9</td>
<td>Introduction to the Exploratory Data Analysis</td>
</tr>
<tr>
<td>10</td>
<td>Introduction to Measuring Air Quality with Sensors and App</td>
</tr>
<tr>
<td>11</td>
<td>Indoor Data Collection Part 1: Around the School</td>
</tr>
<tr>
<td>12</td>
<td>Indoor Data Collection Part 2: Analysis</td>
</tr>
<tr>
<td>13</td>
<td>Planning Your Investigation Part 1</td>
</tr>
<tr>
<td>14</td>
<td>Planning Your Investigation Part 2</td>
</tr>
<tr>
<td>15</td>
<td>Data Collection: Out in the Neighborhood</td>
</tr>
<tr>
<td>16</td>
<td>Outdoor Data Analysis</td>
</tr>
<tr>
<td>17</td>
<td>Using the Exploratory Data Analysis Tool</td>
</tr>
<tr>
<td>18</td>
<td>Citizen Science and Scientific Action</td>
</tr>
<tr>
<td>19</td>
<td>Developing the Project Report</td>
</tr>
<tr>
<td>20</td>
<td>Class Presentations</td>
</tr>
</tbody>
</table>

## The Scenario

There is a rising case of asthma in urban areas in the United States. Researchers have hypothesized that air quality contributes to this phenomenon. However, there may be a number of other causes, including a rise in smoking rates, industrial pollutants, and social stress from a lack of access to life resources. The Town Council of Philadelphia has announced a new community program that will fund projects that are likely to support risk reduction of local asthma cases. These projects may include planting more trees to purify the air, an anti-smoking campaign, and building a park to promote activities that can alleviate stress. The Town Council asks you to research the issue and submit a proposal describing a project that your team would like to fund with evidence from public health and environmental data that supports your proposal.

## Tools

**Exploratory Data Analysis Tool**
- Upload data of interest in csv format
- Outputs univariate and multi-variate graphics

**Air Pollution Visualization**
- PM2.5 and CO AirData estimates across 8 cities in the U.S.
- Crowdsourced pollutant sensor data estimates in Philadelphia

**Gene Expression Microarray Analysis**
- Smokers versus non-smokers microarray data analysis
- Phenotype info, quality control and differential expression

Data visualization web applications
Sample Findings: Teachers had high satisfaction from PD experiences.

### High PD Satisfaction

2019 and 2020 Summer PD Average Teacher Satisfaction

<table>
<thead>
<tr>
<th>Category</th>
<th>Average Rating 19</th>
<th>Average Rating 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
<td>4.39</td>
<td>4.52</td>
</tr>
<tr>
<td>Modules</td>
<td>4.17</td>
<td>4.53</td>
</tr>
<tr>
<td>Materials</td>
<td>4.35</td>
<td>4.50</td>
</tr>
</tbody>
</table>

### Improved application of science knowledge

- Student engaging in authenticity and real-world exploration
- Benefit to student science efficacy beliefs and personal relevance

"...it's a relevant issue that a lot of students will be connected to. It comes with a lot of support from you guys in Penn and it's engaging and the students learn a lot of different kinds of science stuff ...So it's a great project I would totally recommend it."

"I think that these kids...they get geeked up about, 'Okay, let's interpret this data set, so that way we can solve this actual, real world problem.'"
Theme 1: Implementation complexity
- Teachers said that the project had a lot of complex tasks that required them to divide their instructional attention, that often created implementation and classroom management challenges.
- Example quote: “And so having gone through it in actuality, there was a couple things that I found tricky. I found tricky the keeping track of all the devices and the technology. [It was like] ‘Okay. Here. I'll give you this device. Wait a minute. I didn’t…mark who actually had that device.”

Theme 2: Content preparedness
- Teachers said they need more training in bioinformatics content, data literacy, and how to integrate into the biology course.
- Example observation: When teaching a lesson on the human genome from slides put together by bioinformatics scientists during the summer PD, one teacher mentioned that he would be skipping through many of them because he himself was not well-versed on the subject.

Theme 3: Alignment with familiar pedagogical supports
- Teachers wanted more instructional resources that they were familiar with.
- Example quote: “I feel like if I just would have planned this out more, you would’ve had kind of more things to use in our toolbox like handouts, notes, things like that. More substantial things that we can implement.”

Theme 4: Resource navigation and access issues for just-in-time instruction
- Teachers needed easier ways to locate resources such as links to sample datasets at the time of instruction.
- Example quote: “…but I think by the time that I was teaching that myself, it required some more review. When I was [in the PD workshop], I was kind of getting it, but I think [only] because [the bioinformatics instructor] was right there.”
Sample Findings: Students’ learning increased after PBL

Student Learning: Experience Survey

*Cohort 1 students (n=122) demonstrated significant gains in 3 of 5 experience factors measured.

Student Learning: Content Survey

Post survey scores were significantly higher for both open-ended and multiple choice questions. However, the mean scores from the multiple choice questions were relatively low, which indicates that students still had challenges in data literacy.
Research Products: Instructional resources

Asynchronous PD MOOC offered on Edx Platform

PBL curriculum:

1. 22 lesson plans
2. Multiple annotated PowerPoint Presentations
3. Student worksheets
4. Teacher guides

Google Classroom for archived PD resources

Screenshot of a submodule in Edx Edge online course

Screenshot of a Google Classroom
Research Products: Web applications for data visualization

k12bioinformatics.org

- Website provides interactive user-interfaces to teach bioinformatics concepts such as asthma and environmental exposures.
- Consists of three apps:
  1) A data analysis tool for basic data visualization
  2) An app to visualize air pollution measures from low-cost sensors and EPA monitors
  3) An app demonstrating gene expression microarray analyses
- A, B, and C are example features of the air pollution visualization app.

pargasite.org

- Pollution-Associated Risk Geospatial Analysis SITE (PARGASITE) provides visualization of estimate levels of pollutants (e.g., PM 2.5 and CO) in the U.S. for 1997 through 2019 at user-defined geographic locations and time ranges.
- Measures were taken from U.S. Environmental Protection Agency (EPA) regulatory monitor data.
Overall, student interest, engagement and learning increased through project participation.

While teachers enjoyed the summer PD, when implementing with students during the school year, teachers needed more training in bioinformatics, data literacy, and STEM integration strategies.

The curriculum is complex with many moving parts. In order for this exploratory work to be scaled, the curriculum needs to be pared down (this is one of our next steps in the coming year working with project teachers).

During the pivot to online PD, teachers needed more intentional community building structures (we have built this into our next run of the summer PD workshop in July, 2021).
Impact

• Most teacher research for incorporating bioinformatics in high school curricula has been conducted outside of the U.S. This research represents a rare teacher support and PD study on teaching and learning about bioinformatics in a U.S. context.

• The complexity of novel bioinformatics curricula requires greater attention on what teachers need in terms of content and pedagogical content knowledge supports (this is a focus in the coming year).

• This study aims to specify strategies that can support other PD developers with similar aspirations to update high school science curricula to meet important needs in the rapidly expanding field of bioinformatics.

This research was funded by the National Science Foundation (grant #1812738). Any opinions, findings, and conclusions or recommendations expressed in these materials are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.
K-12 Bioinformatics Publications


