GSAF Project Summary
Max Helix

Overview:

This project focuses on developing methods for assessing the outcomes of participating in undergraduate research. There are very few instruments developed for this purpose reported in the literature, and those available have some major limitations and questions regarding validity. A primary goal of this project was to move beyond self-report data to evaluate certain desired outcomes, such as improved content knowledge and understanding of the research process. Our team developed preliminary rubrics to assess these aspects on the basis of poster presentations and other student-generated artifacts. Through pilot testing and further refinement, we hope to create instruments that will be applicable to a broad range of undergraduate research projects.

Assessment team:

Anne Baranger is the director of undergraduate chemistry at UC Berkeley and a chemistry education researcher. Approximately 70% of chemistry majors at Berkeley participate in undergraduate research by directly joining a professor's lab, and all chemistry majors participate in a one semester research project as part of their freshman general chemistry courses. Because it is a significant part of the undergraduate chemistry program, we want to know whether students are receiving the expected benefits from participation in research.

Elisa Stone is the program director for Cal Teach, a program for STEM students at Berkeley that offers a minor in education and the opportunity to complete a California teaching credential. Part of the Cal Teach curriculum involves a one semester research project, and it is in the interest of the program to know if it is achieving its desired outcomes.

Program context and motivation:

Even though a tremendous amount of time and resources are poured into undergraduate research programs at universities throughout the country, their outcomes are rarely assessed in any systematic way. This stands in stark contrast to typical university courses, in which each student is evaluated repeatedly throughout the semester using standardized rubrics. It is widely assumed that there are many beneficial outcomes for students who participate in undergraduate research, but actual evidence for this claim is scarce. A recent review article found that empirical studies in this area are sparse and rely almost exclusively on self-report data, without attempts to validate their results by looking at other artifacts or data sources.[1] While this might be a reasonable strategy for learning about outcomes like "increased confidence as a scientist," it leaves something to be desired if we want to know about things like content knowledge or whether students are thinking critically about their own research process.

Within our assessment team, there is a desire to improve the quality of the research experiences associated with both the Cal Teach program and the chemistry department. However, we currently do not have answers to the following critical questions:

- Are these programs currently meeting their learning goals?
- If changes were made to these programs, how would we determine whether they resulted in improved outcomes?

---

[1] Linn, M.L., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: Impacts and opportunities. *Science*, *347*, 627-632.

In order to address these questions, we need to develop tools for evaluating whether specific desirable outcomes are being achieved by these programs. Although we hope these instruments will be broadly applicable across various STEM disciplines, our initial investigations have focused on the following three groups of UC Berkeley undergraduates:

- Chemistry majors who have joined a research lab
- Students enrolled in Chem 4B. This is the general chemistry course mentioned above that includes a semester-long student-generated research project.
- Students enrolled in UGIS 188. This is the Cal Teach project mentioned above. It is designed to give future STEM educators some experience with the research process.

Purpose and intended use:

We hope that the tools we develop for evaluating the outcomes of undergraduate research will be broadly applicable to various programs and courses. Having instruments like these will allow stakeholders to assess whether they are meeting their goals and where they are falling short. Data collected with these tools will be useful for both making rational changes to the programs and determining whether these changes result in improved outcomes.

Guiding questions:

- What are students learning by participating in these research courses or programs?
- Are students achieving the goals we have for them?
  - What are these goals?
  - Of all the desired goals, which specifically are we looking to assess?
- What artifacts are students already producing that would give us insight into specific learning goals (e.g., content knowledge, ideas about the process of scientific research)?
- What can we learn from this data?
  - Can we create general rubrics to distinguish between students who are at different levels of progress with regard to these learning goals?
  - Will assessment of multiple artifacts from the same student lead to a coherent evaluation of their progress?

Methods and tools:

Defining Learning Goals

Casual conversations with various professors revealed large variations in priorities related to the desired learning outcomes for undergraduate research. In order to determine more broadly whether a research program is "achieving its goals," these goals need to be explicitly stated and agreed upon. An interview protocol is being developed to address this and other issues related to undergraduate research (for current iteration, see Appendix A). To inform the questions about which goals faculty members prioritize, I compiled a list of about 50 outcomes that have been reported in the literature and summarized in Laursen et al. (2010).[2] These were then combined and modified to make a more manageable list of 9 outcomes that faculty will be asked to rank in importance.

Artifact – Responses to Journal Prompts

Students in UGIS 188 are required to submit weekly responses to prompts related to their ongoing research project. An example prompt is the following: "Think about the ways you have analyzed data recently. 1) Describe one example of data analysis you have done. 2) Reflect on a

[2] Laursen, S.L., Hunter, A.-B., Seymour, F., Thiry, H., & Melton, G. (2010). *Undergraduate research in the sciences: Engaging students in real science* (Ch. 2-3, App. D). San Francisco, CA: Jossey-Bass.

situation in which you created a model or explanation based on your data analysis or in which you had to make changes to a model or explanation based on your data analysis." A total of 16 answers to this prompt from a previous semester were gathered for a preliminary analysis. For comparison, an additional 10 responses were elicited from other undergraduate and graduate students at various levels of experience. The research projects described span a broad array of STEM disciplines.

A 6-point holistic rubric was developed (based on Marcia Linn's Knowledge Integration framework[3]) to evaluate the complexity of understanding illustrated by responses to this prompt (Appendix B). In particular, we are looking at how the student conceptualizes data analysis and how it connects to the research process as a whole. Our dataset was evaluated according to this rubric, with good but not great interrater reliability. Future iterations will provide more clarification about how to distinguish between adjacent levels of the scale. Additionally, we may make modifications to the prompt to better elicit the type of data we require.

Artifact – Poster Presentations

All three of our focal student groups are required (or strongly encouraged) to participate in poster sessions at the end of the year where they present their work. This is an inherently reflective task, and student presentations are likely to provide key information about their views on the research process in general, as illustrated by the steps taken in pursuing their own project. Follow-up questions from potential evaluators may also allow us to probe student content knowledge, though this would require the evaluator to be relatively knowledgeable about the science behind each project.

With these goals in mind, a rubric was created to address key factors that we would look for in an advanced/expert poster presentation. Additionally, semi-structured questions were designed to help elicit information about items that were not sufficiently addressed and to evaluate the extent of the student's project-relevant content knowledge. This rubric was pilot tested and modified three times, and the current iteration is attached (Appendix C).

Next steps:

The project is still in its early stages, so there is still a great deal yet to be done. The most immediate priorities include the following:

- Open coding of responses to journal prompts: What are the common themes in student responses? Are there consistent indicators of more or less advanced students, as judged by our holistic rubric (or by other means of assessment)? Can we turn these indicators into an analytic rubric to accompany our holistic one?
- Further development of the poster rubric: Add more nuance to categories that seem too broad, and combine or eliminate categories that are redundant or not informative. Streamline and simplify. Include a more well-defined holistic assessment.
- More rigorous pilot testing of poster rubric: Be slightly more systematic with our approach, and start looking into interrater reliability.

Tips and strategies for others pursuing similar projects:

- Regular meetings are critical for making progress, because everyone has many other projects besides this one vying for attention.
- Planning too much before data collection may be counterproductive, because of how much things will change based on preliminary observations (but obviously don't jump in completely unprepared).

---

[3] Linn, M.C. (2006). The knowledge integration perspective on learning and instruction. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 243–264). New York, NY: Cambridge University Press.

- Be aware: Even “consensus” goals that researchers might have may not overlap completely with those of other stakeholders.

Appendix A

Faculty Interview Protocol

1. How many undergraduates do you usually take on? How do you choose who to work with?
2. What are your overall goals for your undergraduates?
3. Which of the following you consider to be important outcomes for your undergrads? Are there any others not on this list?

- Increased confidence in their ability to do scientific research
- Gaining a deeper conceptual understanding of the field
- Developing their ability to analyze and interpret data
- Developing their ability to propose and design experiments
- Engaging in teamwork, collaboration and communication
- Gain familiarity with what a career in research is like
- Gaining expertise in technical lab skills
- Have an enjoyable experience in the laboratory (push them on this if they select this one - most students enjoy their experience - what else is valuable for students to gain from the experience?)
- Understand what science/research is about (nature of science)

Possible follow up questions: Which one or two goals do you consider the most important? Do you have different goals for different types of students?

4. How do you know that a student is reaching your goals?
    a. What do you look for to show that a student is likely to succeed as a scientific researcher?
    b. How much time do you think it takes working in your lab to see real progress towards [learning goals]?
    c. Do you foster more independent work (students making some of their own decisions rather than asking for all procedures)?
5. Add personal stories of particular undergrads come up. Particularly successful or got a lot out of it or succeeded against other odds (like bad grades).
6. How are undergraduates mentored/trained in your laboratory?
    a. How do you choose (or help the student choose) a project, subproject, or role for a particular student?
    b. What are your expectations for UR work?
        i. number of hours
        ii. semesters of experience
    c. Who mentors them?

        i. Do advanced UR students mentor newer ones, if you have multiple?
        ii. How does a UR student get up to speed on the background of his/her project?
    d. Is there any regularity to graduate student mentoring?
        i. Schedule of meetings with graduate student mentor?
    e. How do you interact with undergraduates?
        i. Do you ever meet individually with your UR students?
    f. Do UR's participate in formal group presentations and meetings
        i. Group and subgroup meetings
            1. attendance?
            2. make presentations?

7. Guidance
    a. How is their work monitored?
    b. How is their thinking about their research project guided?
    c. How are they given help with diagnostics, troubleshooting, problem solving?
    d. Do you feel like your undergrads get a broad knowledge of the field? If so, how does that happen?
    e. What support do undergrads get when they hit the inevitable failures and setbacks of research?
    f. Intervention
        i. When/how/why do you intervene in their projects?
        ii. What do you do if a student isn't getting anywhere in his/her project?
8. What types of (formal or informal) training do you (or your graduate students) provide your UR students in the following areas?:
    a. Finding, reading, discussing, and presenting on the literature. eg. What expectations do you have of your undergrads in terms of learning to read the literature on a regular basis? Do you choose particular papers for them to read?
    b. Lab equipment and techniques
    c. Discipline relevant software and search engines
    d. Keeping good lab notebook
    e. Communication of results
    f. Formal presentations (including Q&A)
9. Typically, graduate students don't get much training to be mentors - do you think it would be helpful?

Appendix B

Knowledge Integration Scoring Guide

Prompt: Think about the ways you have analyzed data recently. (1) Describe one example of data analysis you have done. (2) Reflect on a situation in which you created a model or explanation based on your data analysis.

Main links between independent Science practices (focusing on link #4)

1. Understand the research question of their laboratory and its context in the larger field.

    a. Link question and context

2. Communicate the contribution of their specific project to the larger research question.

    a. Link conclusion/application back to question

3. Design experiments to answer questions relevant to their specific project.

    a. Link experimental design and research questions

4. **Analyze and interpret data in order to construct explanations and models (primary main link).**

    a. Link data analysis to interpretation, AND

    b. Link data analysis/interpretation to explanation/model construction

5. Plan future experiments based on the analysis and interpretation of data and their understanding of the research question.

    a. Link analysis and interpretation to (experimental) research question, OR

    b. Link analysis and interpretation to future experimental design

Complex Links:

Between main link 4 and other links above

Knowledge Integration Level	Score	Criteria
Systemic Link: Elaborates a complex set of connections back to context for and contribution to larger scientific field	**6**	Contains a complex link, and also makes a link between one or more of the following: • research question of laboratory and its context in the larger field • contribution of specific project to larger research question • experimental design to research question relevant to the specific project
Complex Link: Elaborates a connection between data analysis and two or more related concepts	**5**	Contains a full link, and also links data analysis/interpretation to research question and/or future experimental design (eg. both main links 4 and 5)
Full Link: Elaborates a connection between data analysis and explanation and/or model	**4**	Data analysis is clearly described and linked to construction of explanations and/or models. Has clearly answered the data analysis question (eg. within main link 4), but has not linked out to other independent science practices in our list (eg. in main links 1 - 3, or 5)
Partial Link: Data analysis is clearly described, but links to other scientific practices are incomplete or vague	**3**	Data analysis is clearly described, but links to other scientific practices are incomplete or vague Or, experimental design/data collection is clearly described, and analysis tool is identified but outcome of analysis is not Or, data analysis is incomplete, but clear connections are made between components of research process
No link: Responds to question but very surface level	**2**	Individual scientific practices are each incomplete, even if some links are present between practices
Irrelevant	**1**	Did not answer question, or focuses on data collection or instrumentation rather than analysis, or on an analysis technique/tool rather than data analysis, or on big ideas without connecting to specific data
No Information	**0**	No answer/blank

Knowledge Integration Level	Example Responses
6 **Systemic Link**	*[have not identified a response we scored at this level yet]*
5 **Complex Link**	*Recently, I did data analysis on how the thickness of a layer in my solar cell affected its current and efficiency. I'm trying to improve the overall current, so I spent the week trying to find which thickness would be best for outdoor light and indoor light optimization. For the outdoor light, I had an expectation that a thinner layer would work better, and I thought that indoor light would work similarly. However, the indoor light test was actually inconclusive. This struggle, then, is that I am having trouble coming up with a model for how the cell will behave in response to indoor light.*
4 **Full Link**	*I recently investigated the effect of fluoride salts (in combination with differing amounts of base) on product and byproduct formation. I found that although small amounts of fluoride decreases phenol byproduct, the overall product yield decreases slightly.* *Heating of Cu-catalyzed aminations led to a higher amount of phenol byproducts than in room temperature reactions. It was hypothesized that a 0°C reaction might lower this byproduct for thermodynamic accessibility reasons. This was incorrect however, as the phenol content was the same as in room temperature reactions.*
3 **Partial Link**	*I have been working on a series of Li-ion conductive polymers in which I can vary the length of the linker between negatively charged cross-linker nodes and have been using electrochemistry to evaluate and compare their conductive properties.*
2 **No link**	*I used t-tests to analyze data on the effect of mistletoe parasitsm on blue oak water status and then represented it using boxplots, both were done using XLSTAT. It was a fairly small dataset and simple data analysis. From there we built our explanation around the significant data points and past research.*
1 **Irrelevant**	*I don't really do the data processing, that's mostly Tara's job. I do more of the physical work.* *1. I perform kinetic analysis on enzyme that my supervisor purifies. I perform the reaction and allow the enzyme to sit for anywhere from 75 minutes to 24 hours. Once the reaction is done I take the data off the computer in CSV format and do baseline corrections and difference spectra in Excel. I fit the data using Prism. 2. We don't necessarily make models, but I fit the data and extract rates for the reactions I run.*
0 **No Information**	*[blank]*

Appendix C

URE Research Presentation Rubric

Expected Components

Research question

Significance

Experimental Design

Next Steps

Findings/Analysis

Conclusions

Category	Beginning (1)	Proficient (2)	Expert (3)
Can you walk me through your poster from the beginning? [allow presenter to make presentation without interruption]			
Initial Presen-tation	• One or more components absent (eg. experimental design) • Transitions between different components is absent or weak • No reference to background literature • Trivial problem choice, feels like a weak science fair project, not a scientific or broader impact question	• All components present (see above) • Narrates a coherent story, with flow and smooth transitions • Some reference to background literature or broader impact	All criteria of proficient met, *and* • Succinct description of each component (avoids unnecessary detail) • Presentation, and particularly significance, is well-grounded in background literature, what is already known, solves a broader impact question
Communication skills	Lack of confidence, inaudible, hesitation, poor	Inconsistently confident and audible, some	Confident, audible, no hesitations, interacts well with audience,

	interaction with audience, monotone Unenthusiastic about project	hesitations, inconsistent interaction with audience, modulated voice enthusiastic about project	Clearly passionate about research project
Questions	Audience has one or more significant questions for which answers are required for understanding project; low level, asked for clarity Clarifying question that interrupts presentation is needed; starts in middle despite being asked to walk through psoter from beginning	Audience may have questions on details and/or significance, but no further information necessary for basic understanding of project No one present to ask question	High level, asked out of curiosity; pushes presenter to consider new perspectives
Responses to questions	Responds to question poorly, does not address question well	Answers questions adequately eg. provides clarity, but does not provide information 'above and beyond'	Answer question bringing in new perspectives and a wide scope of knowledge
Why is your research and what you learned important? (In other words, if I were to ask you, 'so what?,' how would you answer?)			
Impact	Impact statement does not match project goal or is absent/incoherent	Too general or too specific	Discuss specific significance of their project and how it relates to larger effort What aspects of the current study are novel/adding to the field
[if presenter can ask first question, ask next, etc]. Explain the technique/method/approach you used. Can you say more about _____ ? Why did you use this technique/method/approach? What are the limitations of the study?			

What other technique/method/approach could you have used, if any? (If no other approach appropriate, why not?)			
Content (methods)	Specific content inaccuracies are revealed, or is incoherent	Explain technique/method in a complete but limited way Doesn't do a good job of discussing limitations or alternatives	Explain technique/method and reason for using. Compare to other techniques/methods, outlining advantages and disadvantages.
If you had another month or two to work, what would be your next steps and why? if you had another semester, or year?			
Next steps	Next steps do not match project goals or are absent/incoherent.	Limited to get more data, more variations, larger sample size. Do the same thing again, but more of it. Test reproducibility or only slightly different method or different method with no rationale Too general or too vague It is clear that they did not have the time to do a next step that was already planned, i.e. in middle of project (level 3 cannot be determined) Dead end project, and explanation why, with no description of what could be next	In addition to specific next steps, taking the project to a longer term goal (new application, or connecting to another research goal). Use very different method with good rationale. Rationale for why these are the longer term goals based on data in poster (or elsewhere). Dead end project, clear rationale for why the work will not be pursued , and ideas for a new project or other ways to approach your broader research question

Name:
Research Group:
Project Description (1 sentence):
Holistic Score: _____ (half levels ok, e.g., 2+, 2-, etc.)

How long on project: (___ weeks, ____ months, _____ years)
Planning on continuing on project?
Have you research before as an undergraduate or high school student?

How many times was this poster presented before? (OR, note what time during the poster session the evaluator is talking to student _______)