PROPOSAL OVERVIEW AND OBJECTIVES

Student-centered pedagogies have been developed and demonstrated to improve student learning. For example, in the Process Oriented Guided Inquiry Learning (POGIL) approach, students construct their own knowledge through a learning cycle of exploration, concept invention, and application. Students progress through carefully constructed worksheets in small groups to explore a 'model' (an information rich data display), answer questions that make them think about the model, propose explanations for what they have explored, and then apply those concepts to further problems. As successful as I have found POGIL to be, the assessment strategies I have used are still teacher-centered. Typical programs of high stakes assessments fundamentally undermine my learning goals for students. Rather than focusing on growing as a learner and improving as a scholar, traditional grading incentivizes a grade focus in which students try to maximize the points they can get from any assignment for the minimum of effort, independent of their learning. Poor performance on an assessment is met with the attitude that it is too late to do anything about it, so move on to the next chapter. As I result, I hate grading not because it requires effort, but because students pay no attention to the feedback and do not use the feedback as a learning opportunity. The traditional grading approaches that I have used are a problem because they lead to student frustration, stereotype threat, and attrition from STEM.

A long range goal of my teaching is to help students embrace a life of growth and learning both in the subject matter of Chemistry and in the metacognitive and metaemotional skills they need to succeed beyond the Chemistry classroom. The overall objective for this proposal is to develop a proficiency-based assessment structure for General Chemistry 1 that will be transferrable to other instructors in the General Chemistry program. The *central hypothesis* is that a proficiency-based grading structure will better encourage and motivate students to identify their weaknesses and work to improve. **better a). engagement b). learning outcomes c). attitudes to the subject d). self-efficacy** This hypothesis is based on the demonstrated effectiveness of specifications grading and standards-based grading in other General Chemistry courses.[**? ?**] The *rationale* of this project is that the combination of student-centered pedagogy and student-centered assessment will synergistically improve learning outcomes for students. These course materials will be developed with a team of General Chemistry faculty to maximize the transferability of the materials. Our team is uniquely qualified to successfully execute this project because of our experience implementing POGIL in the large-enrollment General Chemistry classroom.

The specific goals for this project are to:

1. **Develop student learning objectives based and align a proficiency-based grading scheme to those objectives.** We will design a set of learning objectives based on Marzano's taxonomy, which provides a structure to specify appropriate learning objectives at multiple levels of conceptual difficulty. The four tiers of achievement – *Recall*, *Comprehension*, *Analysis*, and *Use* – will be articulated for each learning objective. With this set of learning objectives, we will craft a grading scheme based on a standards-based grading that incorporates both many modes of frequent, low-stakes student work – online adaptive learning systems (Sapling Learning Curve), online homework (Sapling Learning), individual and group quizzes, discussion boards, **other? too much?**.

2. **Design assessments that implement the grading scheme.** For each learning objective and each level of mastery, assignments will be refactored to best align with the stated learning goals. While many materials exist in the form of question banks, rarely are these questions explicitly aligned with learning objectives. This aim will curate materials like quiz questions and build quizzes for each learning unit. In addition, a framework to articulate to students what learning objectives they have mastered and what areas need development will be implemented in the Canvas LMS.

3. **Assess and disseminate the course transformation.** This aim will measure the impact of the transformation on student learning and attitudes. Based on our previous attitudinal data, we will compare the effect of the course on student self-efficacy and engagement, especially measures of student frustration. In addition, we will measure differential impact on students based on demographic factors such as gender. Women tend to underperform men on high-stakes assessments even though they outperform them at low-stakes assessments. We will measure student performance in the transformed course to see if it alleviates this gender bias.

Specific Aim 1 will develop a clear, comprehensive set of learning objectives and an associated grading scheme. *Specific Aim 2* will develop the actual assessment instruments that will assess student learning and communicate to students their progress. Finally, *Specific Aim 3* will measure how effectively the course

transformation has achieved each of its goals and share these results with other faculty.

This project is a *creative* and *original* approach to grading reform in the context of large-enrollment courses. The *expected outcomes* of this project will be course materials – learning objectives, learning assessments, and course infrastructure – whose effectiveness has been tested and is ready to be shared. Finally, this work engages a question very timely question in the broader scholarship of teaching and learning for large-enrollment courses – how do we best build student-centered assessments that match student-centered pedagogies with our limited grading resources?

EXPECTED SIGNIFICANCE

This course transformation has the potential to affect many hundreds of Pitt undergraduates each term. If this transformation proves useful, other faculty in the General Chemistry teaching pool can adopt this approach beginning with the materials we will develop and share. The transformation may also inspire other efforts to reform grading policies in other STEM disciplines through the dB-SERC community. Finally, the results of the course transformation will be disseminated in the national POGIL community.

BACKGROUND AND PRELIMINARY RESULTS

Guided inquiry methods [1 **?** –4] and Process Oriented Guided Inquiry Learning [4] (POGIL) in specific, effectively bring active learning into the science classroom. Since 2012, Garrett-Roe has implemented POGIL in upper-level Physical Chemistry classes and, since 2018, in large-enrollment General Chemistry classes.[**?**] As such, the POGIL pedagogy is already well-established. On the other hand, grading reform is a much newer development and a variety of approaches have been adopted. We here introduce the learning taxonomy that guides our proposed grading scheme, and introduce related grading schemes that have been demonstrated to be efficacious.

A hierarchical taxonomy for student learning objectives

We are basing our proficiency-based grading scheme on a learning taxonomy with a hierarchical structure because that hierarchy more clearly articulates which tasks must be completed before progress to higher levels. This taxonomy is, naturally, very similar to Bloom's taxonomy, but we will highlight how this taxonomy helps make good choices for question design.

Marzano proposed a hierarchical ordering of four cognitive processes – retrieval, comprehension, analysis, and knowledge utilization [5]. The lowest level, retrieval, focuses on the recognition or recall of information. For example, when a student is asked to define a word or provide a synonym, they demonstrate simple declarative level knowledge. The next level, comprehension, includes integrating and symbolizing. The process of integration involves understanding general relationships related to how the information is organized. For example, a student would be able to describe all variables associated with a model and the subsequent relationships between these variables. Symbolizing requires the learner to translate the knowledge into some nonlinguistic or abstract form like a graph or equation. Processes that extend beyond identifying essential and non essential characteristics of a topic fall in the domain of analysis. An analysis task requires the learner to reorganize the information in a way that generates new conclusions. Marzano's Taxonomy proposes five processes associated with the analysis level – matching, classifying, analyzing errors, generalizing, and specifying. Finally, the highest cognitive level, knowledge utilization, includes tasks such as decision making, problem solving, experimenting, and investigating. The key distinction from analysis is the focus of the mental activity on a specific situation rather than the actual knowledge itself. For example, when a student analyzes an error involving a gas law relationship the focus is on the gas law. When the student uses a gas law to make a decision regarding whether or not to transport a container of helium gas, the focus is the situation.

Mastery-based grading

The grading practices of instructors generally fall into two categories – a summative, points-based approach and a formative, mastery based method of evaluation. The three defining characteristics of a mastery based grading system include [6]
- providing a list of content learning objectives,
- awarding the "mastery" of an assignment rather than partial credit, and
- allowing multiple opportunities for students to achieve mastery.

Specific implementations of these principles are found in standards-based grading [7], specifications grading [8], and mastery-based testing [9]. Mastery grading naturally fits with guided inquiry instruction since both are ways of promoting a growth mindset by treating knowledge and learning as an evolving process [10].

Approach

Based on this literature, we propose to revise the assessment model of CHEM 0110 in a mastery-based approach (fig. 1). Each week, students will complete a learning assessment that will have a question for each cognitive level, *recall*, *comprehension*, *analysis*, and *use*. Every three weeks students will have another opportunity to repeat earlier assessments (fig. 1).

In some ways, this approach is very similar to the Quiz, Midterm, and Final model that is commonly implemented. Content units are indicated by color. Within each unit, students take two small assessments (12-15 mins, like a quiz) followed by a longer assessment ($4 \times 15 = 60$ min, like a midterm).

Week		First	Second	Third
			Assessment attempt	
1	Syllabus and introduction			
2	Atomic structure	1.1		
3	Moles and nomenclature	1.2		
4	Chemical Equations and Solutions	1.3	1.1-2	
5	Aqueous reactions	2.1		
6	Enthalpy	2.2		
7	Calorimetry	2.3	2.1-2.2, 1.1-3	
8	Coulomb's Law and Shell Model	3.1		
9	Electron configurations	3.2		
10	Spin and Orbitals	3.3	3.1-2, 2.1-3	
11	Fundamentals of Bonding	4.1		
12	Molecular Shape and MO Theory	4.2		
13	Properties of Molecules	4.3	4.1-2, 3.1-3	
14	Ideal Gas Law	5.1		
15	Intermolecular Forces and Phases			
	Final Exam Period	5.2	5.1	1.1-4.3

Figure 1: The working plan for the assessment schedule. Students have three opportunities to master each learning objective.

Specific aim 1) Develop student learning objectives based and align a proficiency-based grading scheme to those objectives.

Though the object of teaching is for students to learn, the assessments used and the grading scheme applied grading often align with those learning objectives in only an intuitive, unexpressed, or traditional way. Clear-headed assessment of student learning should begin with a clear articulation of the learning objectives that is integrated into all aspects of student work and communication with students. This wholistic approach will help students identify the areas that they have mastered and the areas in which they need to improve.

1.1 Student Learning Objectives

Student learning objectives that span the essential content and the hierarchy of cognitive levels from *recall* to *use* will be refined for the General Chemistry 1 course. Preliminary work has identified 11 major learning objectives and expressed the four mastery levels (Appendix). These objectives will be refined to accommodate consensus among the team of P O G I L instructors at Pitt.

1.2 Grading scheme

Task 1: Levels of mastery Four levels of content mastery are *recall*, *comprehension*, *analysis*, and *use*. Each level will be worth one point for that learning objective on a Knowledge Assessment. We will examine the feasibility of allowing students to only answer *use* questions in Canvas when they have demonstrated master of two or three of the lower levels.

- *Recall* questions test **YYY**. For example, multiple choice, sorting, and multiple selection questions without significant distractors are at the *recall* level. These questions likely can be computer graded. We consider calculations that are algorithmic also to be at the *recall* level. For example,

- *Comprehension* questions test for the ability to **XXX**. Multiple choice, sorting, and multiple selection questions in the presence of significant distractors are at the *comprehension* level and can be computer graded. Selection of the correct answer in the context of significant distractors that express common misconceptions indicates *comprehension*. Additional questions such as explaining and drawing will likely require human graders.

- *Analysis* questions test **ZZZ**. Many free-response calculations fall into this category. Free response numerical result questions can be computer graded. Additional proficiency in the presence of minor errors can be noted by students uploading their work and human graders regrading the response.

- *Use* questions ask students to extend beyond *analysis* and put their knowledge to work for their own goals. *Use* questions involve steps of deciding, choosing, proposing. We anticipate that human grading will be required for all of these student responses.

Examples of questions at each of these levels are provided (Appendix). **What level is a multiple choice calculation with significant distractors? Somewhere we need an estimate of grader time.**

Task 2: Map Progress to Points We need to decide how achievement at different levels maps onto an actual score. **This there a way to do this without taking the plunge? Estimate numbers of students at each level to keep similar grade distribution if students did similar quality work? What would we look at to evaluate if it is ok or needs to change?**

An initial approach is given in table 1.

We anticipate that the existing scheme for translating course points to letter grade will suffice for the course transformation. We can project how performance on the assessments will affect students' grades using typical values for performance in the other categories. Typical students will accumulate 90% of homework, 90% of discussion, 85% percent lab, and 70% on the ACS exam. The expected base score would then be 43%. To pass the course with a C, a target of many biology students, students would need to score 54% on the assessments, which translates to an average of 2.16 on all assessments, which corresponds slightly better than to achieving proficiency in all *recall* and *comprehension* areas. To score an A, a student would need to score 3.6 on all assessments, which corresponds to achieving proficiency in all *recall*, *comprehension*, and *analysis* areas as well as ∼ 60% of *use* questions.

	percent
Knowledge Assessments	55
Lab	15
Homework	15
Discussion board	10
In-class participation	*
ACS Final exam	5
Extra credit	*

Table 1: The revised point structure will be quite similar to the existing point distribution.

Task 3: Build Grading Structure in Canvas **We need to figure out a reasonable way to do this in Canvas. I am not seeing it right now. Options include a) aligning to outcomes, b) quizzes with multiple attempts, c) an assignment group per learning objective, d) modules for each objective e) mastery paths**

Specific aim 2) Design assessments that implement the grading scheme.

2.1 Evaluate, revise, and write assessments

Task 1: Grow a taxonomically organized question bank Existing assessments will first be sorted by their learning outcome (subject area and cognitive level). We have a library of **131** questions that been designed in accordance with this grading scheme for human grading (Shepherd) and ∼ 200 unsorted questions that have been implemented in Canvas for automatic grading (Garrett-Roe). The pool of unsorted questions will be evaluated, grouped according to their learning level, and revised as necessary. Additional questions will be written to fill any identified gaps. We aim to develop question banks of ∼ 6 questions per learning objective (264 questions total).

percent	Grade
93–99	A
90–92	A-
87–89	B+
83–86	B
80–82	B-
77–79	C+
73–76	C
70–72	C-
67–69	D+
63–66	D
60–62	D-
0–59	F

Table 2: We anticipate that the scoring structure will remain unchanged.

2.2 Assemble assessments

All assessments will be drafted for the term based on the item banks developed in (*Study 1*). Each assessment will contain one question, drawn from the item banks, at each cognitive level. Item banks of 6 questions will provide a reasonable likelihood that students will see new questions on each attempt.

2.3 Implement integration canvas

The grade book infrastructure will be developed to calculate grades according to the grading scheme and to efficiently communicate to students their current learning progress. Potential approaches to be explored are quizzes with multiple attempts, multiple assignments with grade pooling, alignment with the Learning Objectives feature, and Mastery Paths in content Modules. We will select, from these options, the

approach that offers the clearest implementation from the instructor perspective and the clearest communication to the student. Appropriate syllabus text will be developed and shared.

2.4 Deploy the course transformation

In Fall 2021, the transformed course will be implemented.

Specific aim 3) Assess and disseminate the course transformation.

some kind of introduction

3.1 Assess Student Learning

Task 1: Course content Compare to 2018-2021 data for assessment items.

Task 2: ACS exam ACS Gen Chem 1 Paired Question exam. (These exams were purchased in 2018 and are available in the Department.) We will compare to both national norms exam data and, perhaps more importantly, to the scores from POGIL sections in 2018 and 2019.

Task 3: Follow on performance Track students going into Gen Chem 2 with other instructors.

Task 4: Student learning self-assessment Student perception of learning gains will be assessed through an instrument developed through `salgsite.org`, which provides easy implementation and analysis of the questionnaires.

The question areas relevant to course learning will include **1.)** how do in-class activities support learning, **2.)** how does assessment structure support learning.

Task 5: Demographic effects We will compare student performance sorted by demographic category and compare against prior years. Gender performance gaps have been identified in STEM classrooms at Pitt and suggested to be present in our Chemistry program. We will identify the magnitude of the gender performance gap in our previous General Chemistry classes 2018–2021 and compare this with student performance after the course transformation.

3.2 Assess Student Attitudes

We will reimplement the surveys of student opinions used in the 2018-2019 terms that captured student. We will also capture data from concurrent POGIL and traditional lecture courses. Instructors will be encouraged to offer students extra credit or some reward for participation.

Task 1: Engagement "I feel engaged during class time." "I do not feel engaged during class time." "I feel engaged when I am working on my homework or out of class assignments." "I do not feel engaged when I am working on my homework or out of class assignments."

Task 2: Self-efficacy The attitude surveys will include questions to assess student perceptions of their self-efficacy in postively coded questions like, "When I face a hard problem, I feel like I can solve it," and negatively coded questions like, "I get frustrated easily when I face a challenging problem." These showed a meaningful shift after the implementation of POGIL in the large-enrollment course compared to traditional lecture.

Task 3: Identity "I am a science kind of person." "I am a biology kind of person." "I am a chemistry kind of person." "I am a physics kind of person." **TODO: review previous results from Paulette.**

Task 4: Anxiety "Taking the assessments (quizzes and exams) was stressful." "I felt relaxed completing assessments (quizzes and exams)." "I am worried about my performance on my next assessment (quiz or exam)." "I feel confident about my next assessment (quiz or exam)."

Task 5: Focus groups In addition to these quantitative surveys, qualitative data will be obtained in focus group sessions. The sessions will elicit feedback on the effectiveness of the assessment approach in the course. Sessions will be voluntary and occur at three points in the semester.

3.3 Disseminate to Other Chemistry Faculty

At the beginning and end of the term of work, the POGIL Chemistry faculty (Madison, Meyer, Garrett-Roe) will meet to review the plan for the transformation and provide input into the project design. When approximately half of the course assessment material has been revised, the POGIL faculty will, as a team, review the status of the project. Feedback will be incorporated. At the end of the project, Madison and Meyer will be able to choose if they will also adopt the class transformation at this point or if Garrett-Roe will run a year as a pilot program. During the summer term, the transformation will be presented to other

faculty in the General Chemistry program.

EXPECTED OUTCOMES

The three aims will together provide a route to demonstrably effective course materials. *Specific Aim* **??** provides the effectiveness of the POGIL approach for this particular course. In that context, *Specific Aim* **??** provides a comprehensive assessment of the strengths and weaknesses of the dynamic course materials I have developed. *Specific Aim* **??** provides a pathway to incorporate the gained understanding into improved materials.

Potential problems and alternative approaches

Though unlikely, it is possible that the instructor of the parallel section of Physical Chemistry 2 will not agree to cooperate. This problem could eliminate the control population. Nevertheless, the diverse portfolio of student feedback I plan to collect is not exclusively dependent on the ability to make the case and control comparison for quiz and exam questions. A meaningful data-set should still be possible given the student self-assessment of learning gains and student feedback on the effectiveness of the dynamic displays to aid their learning.

Another potential problem, is that the student populations in the two courses may indeed be different. Recent experience in the course suggests that both courses have roughly the same proportions of A's and B's, nevertheless, the equivalence tests may fail in this year. Having unequal populations means one cannot say this study proves that POGIL is better in an absolute sense, but that is not the goal of this project. The differences in learning gains and attitude shifts should still be perceptible, and the development and refinement of course materials will still be valuable.

TIMELINE AND BUDGET JUSTIFICATION

The funds from this award will go to dedicated time during summer 2015 to prepare the new simulations and the assessment tools (surveys, questionnaires, rubrics for marking free-response questions, and focus group preparation) (0.5 month summer salary). The project will support two graduate students (1 month effort each). I will mentor the graduate students in evidence-based teaching and assessment methods. The students will then get hands-on experience designing and evaluating the pre- and post-assessments for these interventions and participating in a focus group session (Spring of 2016). An undergraduate programmer will revise and enhance the simulations in Fall 2015 ($10 / hour, 10 hours / week, 15 weeks).

APPENDICES

List of Appendices

1. Example student learning goals for General Chemistry 1 covering both subject area and level of mastery (page 8)

2. Examples of questions at each level of mastery (page 9)

3. Senior staff biosketch (Shepherd) (page 10)

References

[1] Farrell, J. J., Moog, R. S., and Spencer, J. N. "A guided-inquiry general chemistry course" *J. Chem. Educ.* **76**, 570 (1999). doi:10.1021/ed076p570.

[2] Lewis, S. E. and Lewis, J. E. "Departing from lectures: An evaluation of a peer-led guided inquiry alternative" *J. Chem. Educ.* **82**, 135–139 (2005).

[3] Minderhout, V. and Loertscher, J. "Lecture-free biochemistry - A process oriented guided inquiry approach" *Biochem. Mol. Biol. Educ.* **35**, 172–180 (2007). doi:10.1002/bambed.39.

[4] Moog, R. S. and Spencer, J. N. *POGIL: An overview* volume 994 of *ACS Symposium Series* chapter 1, pages 1–13 American Chemical Society (2008). doi:10.1021/bk-2008-0994.ch001.

[5] Marzano, R. J. and Kendall, J. S. *The new taxonomy of educational objectives* Corwin Press (2006).

[6] Kelly, J. S. "Mastering your sales pitch: Selling mastery grading to your students and yourself" *PRIMUS* **30**, 979–994 (2020).

[7] Marzano, R. J. *Formative assessment & standards-based grading* Solution Tree Press (2011).

[8] Nilson, L. B. *Specifications grading: Restoring rigor, motivating students, and saving faculty time* Stylus Publishing, LLC (2015).

[9] B. Collins, J., Harsy, A., Hart, J., Anne Haymaker, K., Hoofnagle, A. M., Kuyper Janssen, M., Stewart Kelly, J., Tyler Mohr, A., and OShaughnessy, J. "Mastery-based testing in undergraduate mathematics courses" *Primus* **29**, 441–460 (2019).

[10] Selbach-Allen, M. E., Greenwald, S. J., Ksir, A. E., and Thomley, J. E. "Raising the bar with standards-based grading" *PRIMUS* **30**, 1110–1126 (2020).

PDF will go here

PDF will go here

PDF will go here