Assessing College Chemistry Laboratory Learning Using Evidence-Centered Design Principles

Donald J. Wink* and Hongyang (Jennifer) Zhang
dwink@uic.edu, hzhan27@uic.edu
Department of Chemistry and Learning Sciences Research Institute, University of Illinois at Chicago

Abstract: This paper presents a study examining student knowledge and skills assessment within a college chemistry laboratory setting. This has been little examined in the learning sciences but principles of evidence-based assessment offer a unique way to design new learning environments and to examine student knowledge. The goals of laboratories are formulated in the form of learning performances with associated tasks and evidence statements that explicitly cover multiple dimensions in an integrated fashion. The learning performances specify the content that students need to know and also the focal practices that need to be used to reason about that content. Specific examples of how the evidence statements provide a basis for coding student knowledge will be discussed with data from a laboratory on acid-base titration.
Keywords: Design; learning; science education

Introduction: The problem of evidence for student laboratory learning

The laboratory learning environment of science instruction is an area of uncertain value in contemporary science (NRC 2007; NRC 2012). This is both because of a lack of empirical data on the value of such settings and the additional cost of teaching laboratory science (Bretz, Fay, Bruck, & Towns 2013). On the other hand, the ability of laboratory courses to engage students in both disciplinary content and the practices of science (Carmel, Herrington, Posey, Ward, Pollock, & Cooper 2019) means that the environment may be a unique location for learning the breadth of science and also providing students with specific experiences that clarify the empirical nature of scientific knowledge. Laboratory is also a potentially rich environment to design instruction and assessments that engage students in more extended and complex investigations (NRC 2005*)*. Most prior work in this environment has focused on discipline-based research perspectives (NRC 2012a). But deeper understanding of the environment—and about student learning—may be available through the specific use of principles developed within the learning sciences. Specifically, the rigorous construction of assessments using evidence-centered design (ECD; Mislevy & Haertel 2006; Pellgrino 2018) offers the opportunity to both improve the quality and meaning of laboratory assessments and to gain deeper insight into student thinking that go beyond the use of grading rubrics and other traditional schemes.

This paper reports research results arising from the use of evidence-centered design in a college chemistry laboratory course. It use the principles of multi-dimensional learning that have been developed in the context of the report *A Framework for K-12 Science Education* (NRC 2012b)*,* which elaborated science learning in terms of what students should know (disciplinary core ideas, DCIs), be able to do (science and engineering practices, SEPs), and understand across science (cross-cutting concepts, CCCs). We investigate the research question: *How do tasks developed using evidence-centered design and associated learning performances and evidence statements specifications permit the study of student mastery of multi-dimensional learning goals?*

Context

This work is done in a general chemistry laboratory class at a large university in the Midwest of the United States that serves a population with no majority ethnic or racial demographic. The university designates a single course for students in traditional chemistry and biochemistry tracks, other academic majors such as biology, and pre-professional tracks, specifically nursing. Students are all enrolled in a related lecture course, but the lecture course does not provide instruction for the laboratory. Hence, the direct instructional materials to support the lab are the laboratory writeup and a short presentation by a graduate teaching assistant.

This implementation of multi-dimensional assessment follows principles and scaffolds associated with K-12 initiatives. In these, disciplinary core ideas, science and engineering practices, and cross cutting concepts are used to specify tasks that contain multiple assessment tasks, presented as learning performances in paper-and-pencil or computer-based tasks (Wink, Gane, Ko, George, Zeller, Goldman, Pellegrino & Kang, 2018; Harris, Krajcik, Pellegrino, & DeBarger 2019). These dimensions are used to a claim about student learning, which is then articulated in the form of evidence statements required to support the claim. The evidence statements also provide a point of specific alignment with the tasks the students are asked to complete. To use this mode of assessment design requires specific constructs that are the basis of learning expectations in the domain (Pellegrino

2018), something that is only recently emerging in higher education science (Stowe & Cooper, 2019). In our case, the DCIs are taken from the *Anchoring Content Concept Map: General Chemistry* of the ACS Exams Institute (Holme, Luxford & Murphy 2015), which are themselves articulated with larger "big ideas" for college chemistry. The other dimensions are drawn from the high school grade band specifications in Appendices F and G of the US *Next Generation Science Standards* (NGSS Lead States 2013). The method of creating learning performances for college science follows the consensus process developed within the General Chemistry Performance Expectations project of the American Chemical Society (Wink, Pazicni & Donovan 2018).

Methods

Figure 1 presents the design and task description of one lab in the new program, the "Analysis of Glass Cleaner" lab. Specific dimensions are combined into two learning performances and then evidence statements outline what students must show in the assessment to satisfy the learning performance outcome. Finally, these are articulated within the lab activity task itself. The lab is implemented over two weeks, with the first week presenting a primarily instructional mode to ensure students have experiences that give them an opportunity to master the knowledge, practice, and concepts required in the evidence statements. In Week 2, students are given a practical assessment where they are responsible for carrying out their own detailed procedure to determine the acid-base parameters (K_a and concentration) of a solution of a weak base. This is done as a lab practical where students receive no additional instruction from the teaching assistant, other than answering technical questions.

<table>
<tr><th colspan="2">Analysis of Glass Cleaners—Dimensions of Student Learning</th></tr>
<tr><td colspan="2">Disciplinary Core Ideas (selected): The laboratory technique of titration serves as a key example for acid–base chemistry and interpretation of titration curves, both conceptually and quantitatively, is an important tool for chemists. (Language from ACCM VIII, G, 1, b).
Science and Engineering Practices: Mathematics and Computational Thinking: Use mathematical computational and/or algorithmic representations of phenomena… to describe and/or support claims and/or explanations (Language from NGSS Appendix F).
Crosscutting Concepts: Patterns—Mathematical representations…and empirical evidence…are needed to identify some patterns (Language from NGSS Appendix G).</td></tr>
<tr><th colspan="2">Learning Performances (Claims)</th></tr>
<tr><td colspan="2">LP 1: Perform titration experiments and represent the pH change with a graph.
LP 2: Interpret titration curves and their patterns with chemical reactions to determine the acid/base strength in terms of dissociation constant K and unknown molarity.</td></tr>
<tr><th>Evidence Statements</th><th>Task Components</th></tr>
<tr><td>ES 1.1 Set up titration apparatus.
ES 1.2 Calibrate and use pH meter.*
ES 1.3 Record data systematically and converting numerical data into graphical representations.
ES 2.1 Describe the titration curve in terms of the change of major species in the solution.
ES 2.2 Locate the equivalence point at the middle of dramatic change area in a titration curve.
ES 2.3 Determine the acid/base strength in terms of dissociation constant K and molarity.

* In Fall, 2019 ES 1.1 and 1.2 were combined.</td><td>Week 1: Task 1. Carry out a strong acid-strong base titration with pH meter; generate and interpret a titration curve.
Task 2. Titrate a weak acid with strong base and determine K_a and molarity of the weak acid based on the titration curve.
Lab Report Questions: Describe a titration curve. Use derivatives to reason equivalence point and buffer region.
Week 2: Lab Practical: Titrate unknown "glass cleaner" solution with strong acid and determine the dissociation constant K and molarity.</td></tr>
</table>

Figure 1. Task / claim / evidence statement alignment of *Analysis of Glass Cleaners.*

This paper's research question concerns the use of the evidence statements in assessment of student knowledge. LP 1 concerns the student performance of the task. This addresses psychomotor outcomes (Bretz et al 2013) and is assessed during the lab and with the student lab notebook. Student learning at a deeper level is assessed in LP 2. In this case, three evidence statements are present: ES 2.1 concerns the molecular contents of the solution; ES 2.2 covers the recognition of where the equivalence point occurs in the pattern of a titration curve; and ES 2.3 relates to the calculation of molarity and dissociation constant. In all cases, the evidence statements were further elaborated into components (not shown). Students were coded as presenting evidence aligned with an evidence statement if all components of the evidence statement were present. The student practical laboratory work that was collected for Week 2 of the experiment was coded for the presence or absence of the evidence statements that support the claim in Learning Performance 2.

Findings

Figure 2 shows the application of the Evidence Statement analysis to one student report from Spring, 2019. In this case, the student shows recognition of the link of chemical species with the titration curve (ES 2.1—indicated in black) but is missing parts of ES 2.2 (purple) and ES 2.3 (red). Hence, although the student presents partial evidence of meeting the learning performance, the full learning performance is not achieved.

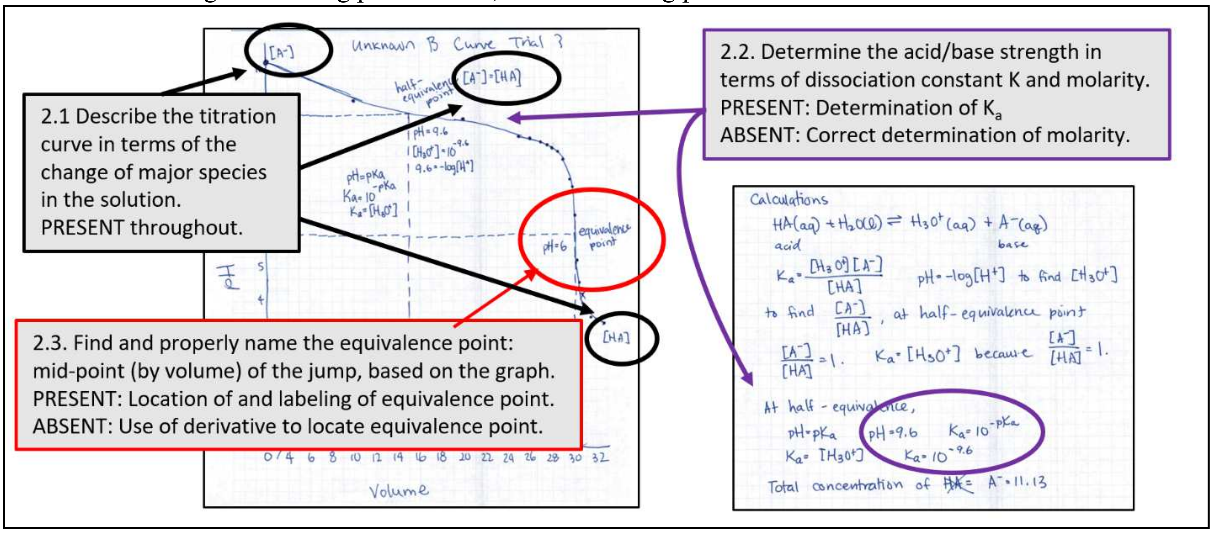

Figure 2. Annotation of student lab report with respect to evidence statements.

The coding system was applied to the laboratory reports of 47 (Spring 2019) and 46 (Fall 2019) students who consented to be in the research study. Students were coded as meeting the learning performance claim requirements if they met the requirement for two of three evidence statements, as is summarized in Table 1.

Table 1: Student results for evidence statements and learning performances for "Analysis of Glass Cleaner" lab

	Spring 2019	Fall 2019
Evidence Statement 1.1	16	44
Evidence Statement 1.2	17	
Evidence Statement 1.3	42	40
Learning Performance 1	20	40

	Spring 2019	Fall 2019
Evidence Statement 2.1	1	3
Evidence Statement 2.2	17	27
Evidence Statement 2.3	14	20
Learning Performance 2	8	17

The results indicate that, in the implementation reported here, many students were able to present evidence matching well the expectation of Learning Performance 1, primarily about the physical tasks associated with collecting and recording data. This included carrying out the mathematical task of representing data with a graph. In the case of Learning Performance 2, only one student actually presented, in the laboratory writeup, evidence of interpreting the different regions of the graph with the different chemical species that are present (Evidence Statement 2.1). A minority of students also interpreted the graph in a way that provided explicit evidence, in the form of a well-reasoned set of calculations, of the practice of mathematical reasoning.

These interpretation of these findings employs the perspective of evidence-centered design, where in this case the evidence allows us insight into both the way the students presented their learning of both content, practices, and cross-cutting concepts. The evidence in this case is that almost all students showed evidence associated with the technically-oriented Learning Performance 1, something that is well-aligned with previous research on the ability of students to match such outcomes (Bretz et al 2013). Far fewer students provided evidence associated with Learning Performance 2, especially in terms of representing the system in terms of the chemical species present. We interpret this, in part, as an indication that students, driven to provide the 'right answer,' are not reasoning in detail but, if they can, are just adopting algorithmic strategies. It also likely reveals issues with how well the students are linking this laboratory work to larger conceptual questions such as equilibrium.

These results indicate that students are doing well in the stated learning performance goals of carrying out titrations and determining "acid-base strength" as a numerical property. However, there are implicit conceptual goals of these as well, such as relating their work to big ideas of equilibrium and chemical reactivity that are not demonstrated in the performance task. These results provide insight, beyond what could be done with conventional grading, of the gap between psychomotor and conceptual learning outcomes (Bretz et al 2013).

Conclusions

Our results in this initial use of ECD in the college chemistry laboratory indicates that the structure of the evidence statements and learning performances can be used to interrogate student work with respect to evidence of learning in multi-dimensional settings. It also reveals the challenges, which are not unique to this project, of engaging students in a way that they provide evidence of the use of practices in particular. It also demonstrates the promise of the use of deeper models for assessment design, as developed within the learning sciences, for use in novel settings that transcend the information available in typical rubrics that only look for completion of tasks. However, as with other translational research, additional iterations will be needed to provide a firmer basis for reliable conclusions about student learning.

References

AAMC (2019). *AAMC-HHMI Scientific Foundations for Future Physicians*. Washington, DC: Association of American Medical Colleges.

Bretz, S. L., Fay, M., Bruck, L. B., & Towns, M. H. (2013). What faculty interviews reveal about meaningful learning in the undergraduate chemistry laboratory. *J. Chem. Educ.*, *90*, 281-288.

Carmel, J. H., Herrington, D. H., Posey, L. A., Ward, J. S., Pollock, A. M., & Cooper, M. M. (2019). Helping students to "do science": Characterizing scientific practices in general chemistry laboratory curricula. *J. Chem. Educ., 96(3*), 423-434.

Hofstein, A. & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*, 28-54.

Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & DeBarger, A. H. (2019). Designing knowledge-in-use assessments to promote deeper learning. *Educational Measurement: Issues and ractice 38*(*2)*, 53–67.

Holme, T., Luxford, C., & Murphy, K. (2015). Updating the general chemistry anchoring concepts content map. *J. Chem. Educ. 96*, 423-434

Mislevy, R., & Haertel, G. (2006). Implications of evidence-centered design for educational testing. E*ducational Measurement: Issues and Practice*, *25(4)*, 6–20.

National Research Council. (2005). *America's Lab Report: Investigations in High School Science.* Washington, D.C.: National Academies Press.

National Research Council. (2012a). *Discipline-based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering.* Washington, D.C.: National Academies Press.

National Research Council. (2012b). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, D.C.: National Academies Press.

NGSS Lead States. (2013). *Next Generation Science Standards*. Washington, DC: National Academies Press.

Pellegrino, J. W. (2018). Assessment of and for Learning. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman & P. Reimann (Eds.), *International Handbook of the Learning Sciences*, New York: Routledge. pp. 410-421.

Stowe, R. L. & Cooper, M. M. (2019). Assessment in chemistry education. *Israel Journal of Chemistry*, 59, 598-607.

Underwood, S. M., Posey, L. A., Herrington, D. G., Carmel, J. H., & Cooper, M. M. (2018). Adapting assessment tasks to support three-dimensional learning. *J. Chem. Educ.*, *95*, 207-217.

Wink, D. J., Gane, B. D., Ko, M-L., George, M., Zeller, L., Goldman, S., Pellegrino, J., Kang, R. (2018) Developing interdisciplinary competencies for science teaching and learning: A teacher-researcher professional learning community." In *Rethinking Learning in the Digital Age. Making the Learning Sciences Count*, The International Conference of the Learning Sciences , Volume 3 Kay, J. & Luckin, R., eds, 1521-1522.

Wink, D. J., Pazicni, S., & Donovan, A. (2018) The ACS General Chemistry Performance Expectations (GCPE) program: From task force to community. Paper presented at 25th Biennial Conference on Chemical Education, Notre Dame, IN.

Acknowledgments

We acknowledge the support of the Single Step Foundation and contextual support of the American Chemical Society General Chemistry Performance Expectations program for this work.