

INSTRUCTORS' FORMATIVE ASSESSMENT IN UNDERGRADUATE BIOLOGY: INFLUENCES, CONTEXT, AND PRACTICES

Authors: Julie A. Birt, Mojtaba Khajeloo, Marcelle A. Siegel, Elizabeth M. Gammel, Hai T. Nguyen, Linh T. Ngo, Keala Cummings and Bethany R. Mordhorst, University of Missouri

Introduction

Formative assessment includes all the activities undertaken by both teachers and students that provide informative feedback to them so that they can modify their teaching and learning activities (Black and Wiliam, 1998). Based on educational studies,urpso the use of formative assessment systems can improve students' performance in higher education (Fuller, 2017; Hattie and Timperley, 2007; López-Pastor and Sicilia-Camacho, 2017; Witzig, Freyermuth, Siegel, Izci and Pires, 2017). Increased use of formative assessment strategies in undergraduate biology courses is associated with gains in both student content learning and student attitudes about biology (Connell, Donovan, Chambers, 2016). While the benefits and difficulties of formative assessment have been studied extensively in K-12 education (e.g., Gottheiner and Siegel, 2012; Ruiz-Primo and Furtak, 2007; Siegel and Wissehr, 2011), less evidence exists at the undergraduate level. Despite the acknowledged significance of formative assessment, there is still a pressing need in higher education for instructors to intentionally plan to include formative assessment and use it (Minbirole, 2016).

In this study, drawing on the personal practice assessment theory (PPAT) model from Box, Skoog and Dabbs (2015), we investigated college instructors' theories and practices about formative assessment. The PPAT model addresses what and how instructors assess and what contextual elements are influential (Box et al., 2015). The components of PPAT framework (Figure 1), are used here to better understand instructors' decision making within their contextual conditions for enacting formative assessment.

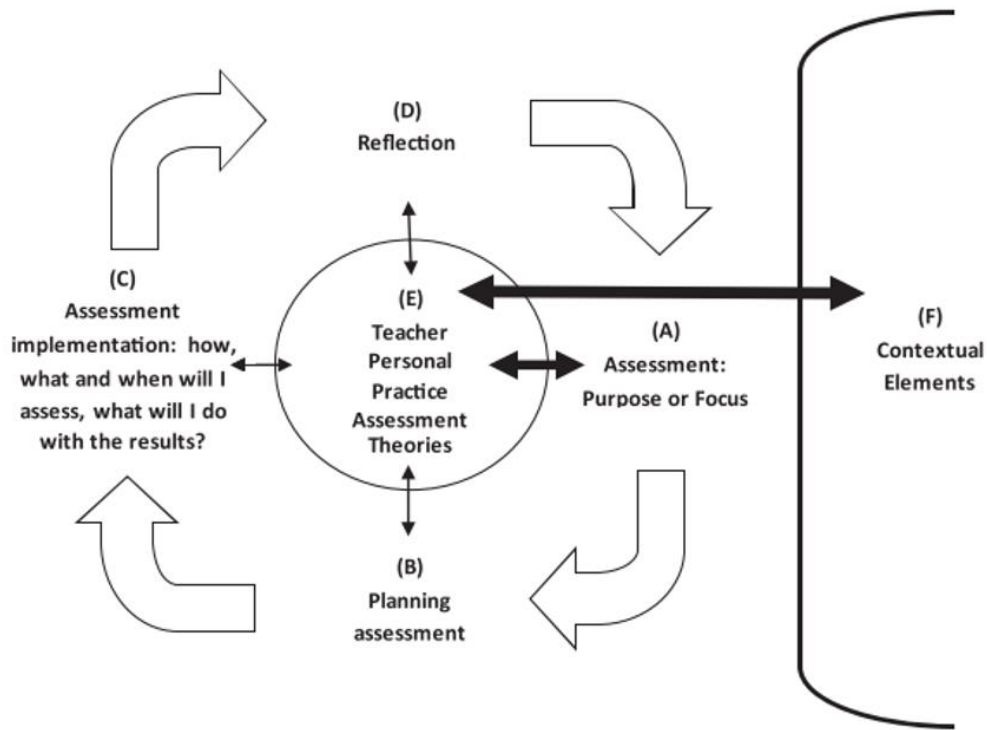


Figure 1: Personal Practice Assessment Theory Model as represented by Box, Skoogs, and Dabbs, 2015.

Conceptual Framework

At the center of the PPAT model is the instructor’s personal practice theories. Personal practice theories are a “systematic set of beliefs [theories] which guide the instructor and come from prior life experiences [personal] and classroom experiences [practical].” (Cornett, Yeotis, and Terwillinger, 1990, p150). Adapted to formative assessment this definition includes an instructor’s theories about formative assessment which come from their previous personal and practical experiences. These PPTs are often not explicitly stated by the instructor and yet they are easily defined by instructors when prompted to consider their purposes and intentions in the classroom (Clandinin, 1986). Few studies have given prominence to instructors’ unique PPTs as they relate to assessment. Box et al. (2015) proposed “that instructors’ PPATs influence what

and how they assess and are influenced by contextual elements, both internally constructed and externally imposed.”

We have adapted the PPAT model to further hone in on instructor’s design and use of formative assessment in their college science classrooms. In this model the instructors PPAT (E) influence their beliefs about the purpose and focus of formative assessment (A). This includes things such as guiding future instruction or assessing students current understanding, and is influenced by instructors’ reflections (D) on previous use of formative assessments and contextual factors (F). The contextual factors can include both internal and external factors that act as barriers or facilitators of the instructors desired assessment practices. For instance, instructors’ experience of assessment, either from student experience or personal practice, can act as an internal contextual factor on their assessment decision makings. Likewise, class size and time availability, as external contextual factors, affect implementation of a particular assessment practice (Levin, He and Allen, 2013; Sweeney, 2003).

In this model, the instructors’ planning of formative assessments (B) is influenced by the PPATs (E) as well as their understanding of the purposes and focus of formative assessment. They must decide the amount of class time to dedicate to formative assessment and how students should benefit from completion or performance. This implementation (C) is influenced by the instructor’s plans (B) as well as their PPATS (E). Importantly, continuing in a cyclical motion, the instructors then reflect on the outcome (D) of the formative assessment process and alter their PPATs (E) and their views of the purposes and focus of formative assessment (A). PPTs can change over time as instructors gain more teaching experience and as their teaching contexts change (Levin et al., 2013; Sweeney, 2003; Cornett, 1990).

Categorizing an instructor's PPATs in the bounded context of their current teaching situation is a helpful way to visualize their reasoning for formative assessment choices. They can also provide an understanding unique contextual elements that constrain or promote the achievement of the desired practices to align with their theories. Thus, the purpose of our study is to define and analyze two experienced college instructors PPATs as they relate to formative assessment in their classrooms while identifying any contextual factors which might influence enactment of these PPATs. Our research questions included:

- 1) What personal practice assessment theories influence the implementation of formative assessment by our college instructor participants?
- 2) What contextual elements constrain or facilitate the use of formative assessment by our college instructor participants?

Methods

Research Design

We investigated the complex and dynamic environment of instructors' practice of formative assessment in context using the "multiple-case study" approach (Yin, 1994). Two experienced biology instructors currently teaching an upper-level biology course with differing backgrounds and roles at the university were carefully selected as participants for this study. Tasha is a teaching professor in biological science and at the time of study was the instructor of a junior-level evolution and ecology course. Jack is a professor in biological science teaching senior-level behavioral biology. The cases were bounded within the context of a single, upper level biology course that the participant instructors were teaching during the semester of study.

Data Collection

Data collected included 3 or more observations of classroom practices for each instructor. To strengthen the data, observations were videotaped, audio recorded and included field notes from at least 2 of the researchers. Observations occurred within a single module in the course to capture how the instructor implemented assessment throughout. Detailed field notes were collected to record all instances of planned or spontaneous formative assessment and student response as observed by the researchers.

Another primary source of data include background interviews with each instructor to discuss their teaching philosophy. These interviews also serve to allow the instructors to explain their planning and reasoning for the observation module. Following the observations, another interview was conducted to allow for targeted questioning of the instructor about any instances of formative assessment noted during the observation module. Instructors were asked about how they interpreted the students' responses and what they planned to do with the information.

Student focus group data were also collected to assess the instructors' formative assessment practices from the students' point of view. Following the observation modules, students were asked to reflect on the types of assessment they had experienced throughout the class and how that affected their learning.

Secondary data sources include artifact collection including written assessments and the syllabus for the course.

Data Analysis

Each instructor's case was analyzed individually to categorize each of the 6 components of their unique PPAT framework. Primary analysis included open coding of the field notes to determine all implementations of formative assessment used by the instructor. To reach a

consensus about this coding, the research team collectively reviewed key pieces of evidence within the classroom videotape data and field notes. The student focus group (SFG) data also proved to be a rich source of information about student interpretation and perception of the instructor's assessment practices. The research group worked together to analyze the data and met regularly to reach consensus as to the conclusions around each of the components. To define the instructors PPATs and beliefs about the purpose and plan for assessment, at least two researchers independently coded each of the instructor interviews and met with the research team to reach a consensus. Finally, to address research question 2, in reference to contextual influences on their formative assessment practice, the instructor interviews were analyzed and reviewed by at least two researchers. Contextual influences were either characterized as barriers or as facilitators to implementing formative assessment.

Results

Tasha - Carefully planned assessment for student learning and motivation

Tasha is an experienced teaching professor whose role at the university primarily consists of teaching in that she teaches several biology courses each semester. During the time of study, she was the instructor of a junior-level evolution and ecology course with approximately 125 students. The module in which her practices were observed was that of the origin and history of life unit toward the midpoint of the semester.

Tasha's PPATs were best characterized into two statements in that she based her assessment decisions upon a belief that:

- 1) Assessment is a learning opportunity for students and a chance for students to see their own understanding rather than memorization

2) Assessment is a strong motivational tool for students.

This is most clearly evidenced in her discussion surrounding the purpose of assessments. She expressed several purposes for including assessment in her highly structured course. Tasha thought assessment was important for helping students learn the concepts in the course as well as a chance to reward participation. In the interview she says that “It helps the students learn the concepts from the course but it also is an incentive for them to put in the effort to do so. One other thing, it is also very useful for me to know what they are not getting, what they are not understanding. That is something that is very valuable about assessment.” Her students appreciated these frequent assessments as motivation to attend the lecture rather than as increasing conceptual learning. Tasha also stated that the assessments were planned to engage the students in the lectures, allowing them to stop taking notes for a moment to answer a question, in order to hold their attention better.

When planning her assessments, Tasha used several different formats to address student motivation and learning through assessment. She added bi-weekly exams which only took the first half of a single class period. She felt these frequent exams allowed students to have smaller amounts of material to study and lowered stress levels for the students taking the course. The students liked the frequent testing because it meant more points in the course and a chance to drop their lowest exam score. Her exams in the evolution and ecology course included essay questions to test student understanding and higher-order thinking skills. During class, Tasha also planned clicker questions using the Top-Hat software which asked students to respond to a multiple choice question to “confirm that they got key points”. Top-Hat questions were often old exam questions from previous semesters. Tasha also had students frequently turn in short,

written answers to application questions covered in the course readings or previous lectures.

Finally, she developed reading questions in the online learning management system, Canvas, that the students were required to answer to help guide them in their reading.

Tasha: Assessment Development Model

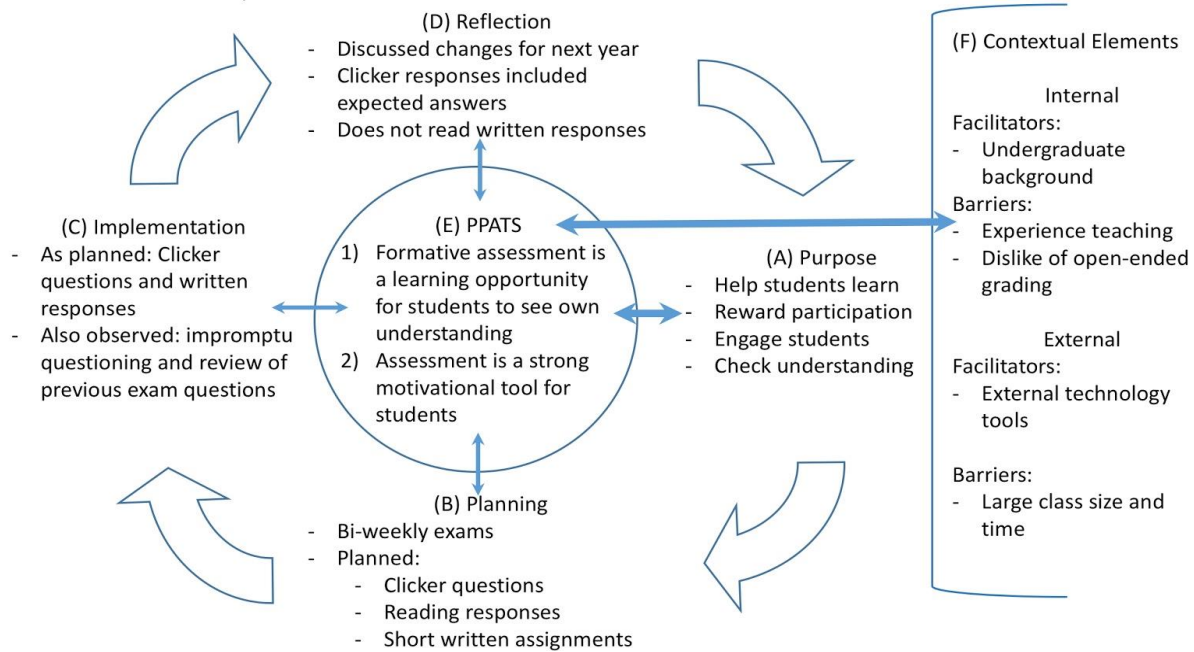


Figure 2. Tasha’s Assessment Development Model

Tasha’s implementation of assessment included the previously mentioned clicker questions and written answers of deeper thinking questions. We also recorded frequent, direct questioning of students about their understanding as well as a review of several highly missed exam question on a previous exam. When reviewing questions from the previous exam, Tasha included exam data to demonstrate the numbers of students who had correctly answered the question. This directly aligns with her PPAT related to using assessment opportunities as a motivational tool and gauge for understanding.

Upon reflection of the results of her assessment practices, Tasha often talked about how she might change the course next year rather than modifying for the current semester. She states, “so that tells me when I am teaching this material next time, I need to look more carefully and think more about how to teach that concept in a way, because I thought that this question was not difficult, I realize it was harder than I thought it [was], I also am thinking about ways I can enhance the likelihood that they would be able to answer such question.” She also mentioned that her past experiences with this course allowed her to anticipate student problems and to keep to her plan to get through the materials. When discussing some complex answers entered into the clicker system during observation, she said, “Occasionally, it will lead to a little diversion, but we get back to what I planned. I have to admit that it’s not too often that I get a student who says something that I haven’t heard before or wasn’t expecting.” In reference to the use of the short, written assessments, Tasha does not read the responses but rather, upon reflection, admitted that they were included as a learning and motivational tool for students to do the background reading instead of a chance for her to gauge student understanding. This clearly aligns with her PPATs which put the focus of assessment on the students’ learning and motivation, but in no way influences her future practices.

Our analyses revealed several contextual elements which facilitate Tasha’s use of formative assessment the most influential of which is that of Tasha’s background as a dual biology and psychology major during her undergraduate years and her experiences teaching. Tasha described that her undergraduate experience allows her to see that there are other ways to gauge student understanding such as the essay questions used in her psychology courses. She also described how technology like Top-Hat and the learning management system helped her to implement more formative assessment.

Conversely, we found that Tasha's many years of teaching experience also acted as a barrier to her formative assessment practice in that she felt that she already knew what to expect from students and did not often allow herself time or opportunity to consider that students might have different understandings than those she has previously experienced. Tasha also expressed a dislike of grading open-ended questions and that she would prefer to devote her limited time to writing good multiple choice questions rather than grading. Finally, a large class size and limited time acted as barriers to implementation of other assessment practices.

Jack - Assessment as a diverse and stress-free student learning experiment

Jack is an experienced professor whose role at the university consists of both teaching and research responsibilities. He was the instructor of a senior-level behavioral biology course with approximately 65 students during the time of the study. The module in which his practices were observed was that of the sociality unit which covered collective movement and group living toward the midpoint of the semester.

Jack's PPATs were best characterized into two statements in that he based his assessment decisions upon a belief that:

- 1) assessment should be a stress-free learning experience for students and
- 2) assessment should allow for diverse rather than fixed answers and include students' interests.

Jack: Assessment Development Model

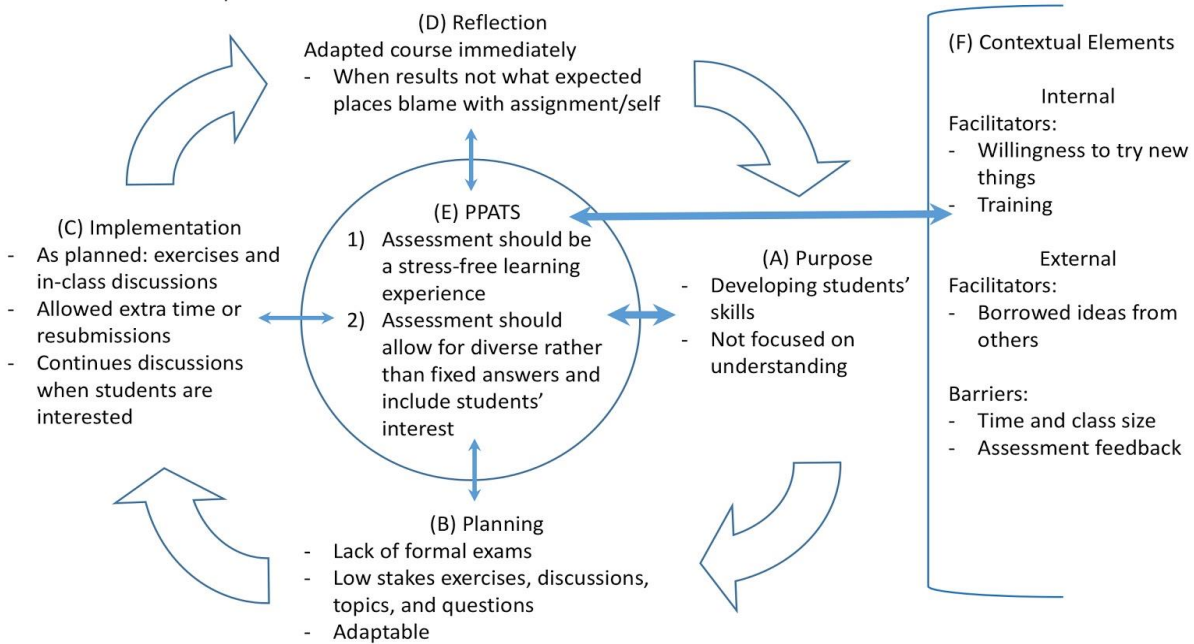


Figure 3. Jack's Assessment Development Model

During his interview he explained this: “partly I find exams unpleasant, but partly they seem very constraining, for everybody is supposed to answer the same thing the same way. But if they’re really getting engaged with it, and thinking about how this relates to their lives and what they learned, that’s got to be more interesting for them.” This is clearly evident in his purpose of assessments that focus on developing student skills rather than understanding of content. Jack stated “but I want them to go away with this, I want to change the way they see the world around them, the way they perceive other species of animals, and make them realize how incredibly amazing other creatures are.” One student in the focus group noticed that “He wants it to be about what you're doing and what you're learning; it’s not about necessarily looking for the right answer.”

When planning his assessments, Jack used several different formats in order to address his goals of focusing on diverse skills and facilitating a stress-free environment. From his

syllabus Jack explains the lack of formal exams: “One of the best-documented findings in recent educational research is that people incorporate information best when they have opportunities to recall it soon after learning it, and when they have the opportunity to explore and apply what they’ve learned in a social setting. So instead of testing you on the information six weeks later, I will provide low-stakes exercises, questions, and discussion topics. You will also keep a course log, a short recap of two main points from each class period.” When discussing planning, Jack often notes his adaptability. For example, the course log is an “evolving concept” and he is flexible with the final project.

Jack’s implementation of assessment included in-class discussion activities and modeling exercises in either small groups or with the whole class. Students were asked to read an article or watch a video the night before and then discuss it in relation to another scenario. Jack often adjusted his assessments showing that his implementation of planned assessments is best characterized as adaptive and strategic. For example, the course log turned into a superficial record of the activities in class so he modified it to ask students to reflect about “what did they get out of it? What do you think differently about now? What did you learn that was interesting to you, because I’m not testing you on this material.” His PPAT related to student interests and diversity are also evident in his willingness to continue on a topic when students were interested. One student notes “And sometimes I think we spend more time talking about certain topics more than others, and I think a 1-day topic kind of gets pushed into a 2-day topic, I think things get pushed back because he's trying to keep what our interests are as his first priority.” In keeping assessments stress-free he also often allowed students extra time or resubmissions to improve their work.

Upon reflection, Jack showed changes in his immediate assessment practices with the course logs, resubmissions, and changes in long-term assessment practices. When discussing the class project, Jack states “The project generates a certain amount of anxiety because partly I’m always trying to change how I do it as I go along and change it depending on what happens in class.” When reflecting on the in class writings, he was not satisfied with the quality of the writing and states “Those in class writings are not as effective as they could be, and I’m not really sure exactly what to do about those. I think they could be better.” Often Jack seemed unsatisfied with his assessment practices and reflected on various ways to modify it.

Our analyses revealed several contextual elements which facilitate his use of formative assessment to find new modifications to use in his classrooms. He often borrowed ideas and assessment strategies from other instructors and instructors. In regards to internal factors, Jack’s willingness and ability to try new assessment strategies also acted to facilitate his PPATs. In regards to the high percentage of As and Bs in his course, he asserted, “I’m a full professor so I don’t care what people think-I would like everyone to get an A in the class.” Therefore, Jack’s job security and his enjoyment of reading student writing influenced his PPATs and assessment practices.

Alternately we found that Jack was constrained in his assessment practices through three main barriers: time, training and class size. Jack was constrained for time to plan future assessments or give feedback on past assessments as well as complete his research responsibilities. He also lacked time to keep updated on the teaching journals. In his interview, Jack often cited his lack of instructor training when discussing his assessment practices and class size. The original version of the course was designed from a 30-student class but at the time of the study he had about 65 students in the course. For him, this affected how successful the class

discussions were and how quickly he could return written assignment feedback. Overall it all comes back to time constraints.

Cross-case analysis

When the two cases were compared we found several similarities among the instructors PPATs and practices. Both instructors emphasized that their formative assessments were planned to facilitate student learning and understanding of the course content rather than rote memorization. Tasha used in-class writings to make students consider application questions while Jack used class discussions of new scenarios to promote learning. The two instructors also drew on individual and group work and included individual writing responses and group activities. While Tasha used this as a motivational tool to show students that the content goals are achievable, Jack used his class discussions to promote a diversity of responses and interaction within the class.

The instructors enacted diverse formative assessment techniques ranging from impromptu questioning (Tasha) to written journals and class discussions (Jack). Instructors' different backgrounds lead to different internally constructed beliefs and in turn influence their PPATs as also seen by Box et al. (2015). Both instructors used assessment strategies they enjoyed. Jack enjoys reading student writing while Tasha would rather spend her time writing challenging multiple choice questions instead of reading written responses. Jack's tenured status allowed him some confidence in making immediate changes to his course while Tasha did not have this confidence and rather made minor changes to her course in subsequent semesters. It was interesting to note that when students were not performing on an assessment as expected, Jack often attributed this to how he designed the assessment while Tasha instead put the burden

of performance on the students' lack of understanding or ability to answer a certain type of question. While the two instructors' practices and PPATs represented a broad range of practices, in the end, they were both primarily concerned with student learning.

Discussion

There is a lack of observed examples of formative assessment use in college science classrooms (D'Avanzo et al., 2012) and this study seeks to fill that gap by presenting a focused glimpse into the private classrooms of college biology instructors. The detailed cycle of assessment development presented for each of the two instructors illustrates a variety of reasoning used by college biology instructors. It also considers the varied findings about the college instructor's practices in each of their unique internal and external contexts. This study also raises questions about how an instructor's PPATs might be affected or formed by student course evaluations as this was not considered when discussing the instructor's previous teaching experiences.

Finally, this study supports the conclusions of other researchers, such as Box et al. (2015), that an instructor's PPATs affect how they understand the purpose of formative assessment and how they navigate the barriers to their assessment practices. We also found that the contextual barriers could be influential to practice and in turn influence the instructor's PPATs (eg. Tasha's experience teaching prevented her from considering other types of possible responses to her assessments). In a similar manner, Levin et al., (2013) found that contextual elements were as influential on instructor practice as the instructors' PPTs. While this study begins to address a gap in the research literature, our findings show the need to clarify how contextual barriers might affect assessment practices and instructor PPATs. It would be

interesting to continue to study how these instructors' PPATs might differ when considered in the context of teaching a different course level or topic as was seen in Sweeney (2003) in which a beginning instructor's personal practical theories changed with a new teaching context.

In conclusion, our findings provide practical implications for university instructors and professional developers in terms of making visible the reasoning and practices that affect formative assessment. While contextual barriers to implementing more formative assessment depending on the instructor, we found that overall, if the instructors can find an assessment strategy that not only aligns with their PPATs but also is enjoyable they will work to implement it in their classrooms.

References:

- Black, P., and Wiliam, D. 1998. Assessment and classroom learning. *Assessment in Education: principles, policy and practice*, 5(1), 7-74.
- Box, C., Skoog, G., and Dabbs, J. M. 2015. A case study of teacher personal practice assessment theories and complexities of implementing formative assessment. *American Educational Research Journal*, 52(5), 956-983. doi: 10.3102/0002831215587754
- Brownell SE, and Tanner KD 2012. Barriers to faculty pedagogical change: lack of training, time, incentives, and ... tensions with professional identity. *CBE Life Sci Educ* 2012;11:339-346. doi: 10.1187/cbe.12-09-0163
- Clandinin, D.J. (1986) *Classroom Practice: Teacher Images in Action*, London: Falmer Press
- Connell, G. L., Donovan, D. A., and Chambers, T. G. 2016. Increasing the use of student-centered pedagogies from moderate to high improves student learning and attitudes about biology. *CBE Life Sciences Education*, 15(1), 1–15. doi: 10.1187/cbe.15-03-0062
- Cornett, J. W., Yeotis, C., and Terwilliger, L. 1990. Teacher personal practice theories and their influences upon teacher curricular and instructional actions: A case study of a secondary science teacher. *Science Education*, 74(5), 517–529.
- D'Avanzo, C., Anderson, C.W., Hartley, L.M., and Pelaez, N. 2012. A faculty-development model for transforming introductory biology and ecology courses. *BioScience*, 62 (4), 416-427. doi: 10.1525/bio.2012.62.4.12
- Fuller, K. 2017. Beyond Reflection: Using ePortfolios for Formative Assessment to Improve Student Engagement in Non-Majors Introductory Science. *American Biology Teacher*, 79 (6), 442-449
- Gottheiner, D.G., and Siegel, M.A. 2012. Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education*, 23, 531-557.

- Hattie, J., and Timperley, H. 2007. The power of feedback. *Review of educational research*, 77(1), 81-112.
- Levin, B. B., He, Y., and Allen, M. H. 2013. Teacher Beliefs in Action: A Cross-Sectional, Longitudinal Follow-Up Study of Teachers' Personal Practical Theories. *Teacher Educator*, 48(3), 201–217. <https://doi.org/10.1080/08878730.2013.796029>
- López-Pastor, V., and Sicilia-Camacho, A. 2017. Formative and shared assessment in higher education. Lessons learned and challenges for the future. *Assessment and Evaluation in Higher Education*, 42(1), 77-97. doi: 10.1080/02602938.2015.1083535
- Minbiole, J. 2016. Improving Course Coherence & Assessment Rigor: “Understanding by Design” in a Nonmajors Biology Course. *American Biology Teacher*, 78(6), 463–470. <https://doi.org/10.1525/abt.2016.78.6.463>
- Ruiz-Primo, M. A., and Furtak, E. M. 2007. Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of research in science teaching*, 44(1), 57-84.
- Siegel, M.A., and Wissehr, C. 2011. Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, 22 (4), 371-391.
- Sweeney, A. E. 2003. Articulating the Relationships Between Theory and Practice in Science Teaching : A model for teacher professional development. *Teachers and Teaching: Theory and Practice*, 9(2), 107–131. <https://doi.org/10.1080/1354060032000089441>
- Witzig, S.B., Freyermuth, S.K., Siegel, M.A., Izci, K., and Pires, J.C. 2013. Is DNA alive? A study of conceptual change through targeted instruction. *Research in Science Education*, 43, 1361-1375.
- Yin, R.K. 1994. *Case Study Research: Design and methods*. Thousand Oaks, CA: Sage Publications, Inc.