5-15-2020

Science Teachers' Understanding of Formative Assessment and its Practice with Three-dimensional Teaching and Learning

Yotah Koulagna

ACCEPTANCE

This dissertation, SCIENCE TEACHERS' UNDERSTANDING OF FORMATIVE ASSESSMENT AND ITS PRACTICE WITH THREE-DIMENSIONAL TEACHING AND LEARNING, by YOTAH KOULAGNA, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the College of Education & Human Development, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chairperson, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty.

____

Renee S. Schwartz, Ph.D.
Committee Chair

____

Patrick J. Enderle, Ph.D.
Committee Member

____

Jodi Kaufmann, Ph.D.
Committee Member

____

Gertrude Tinker Sachs, Ph.D.
Committee Member

____

Date

____

Gertrude Tinker Sachs, Ph.D.
Chairperson, Department of Middle and Secondary Education

____

Paul A. Alberto, Ph.D.
Dean
College of Education & Human Development

AUTHOR'S STATEMENT

YOTAH KOULAGNA

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulations prescribed by the author in the preceding statement. The author of this dissertation is:

Yotah JAB Koulagna
6478 Dunmoor Drive
Jonesboro, GA 30236

The director of this dissertation is:

Renee' S. Schwartz
Department of Middle and Secondary Education
College of Education
Georgia State University
Atlanta, GA 30303

CURRICULUM VITAE

Yotah Koulagna

ADDRESS: 6478 Dunmoor Drive
Jonesboro, GA 30236

EDUCATION:

Ph.D.	2019	Georgia State University Teaching and Learning
Masters Degree	2010	Western Governors University Science Education
Bachelors Degree	2001	Clayton State University Health Administration
Bachelors Degree	1992	University of Yaounde Biology

PROFESSIONAL EXPERIENCE:

2008- Present	Science Teacher Love Joy High School 1578 McDonough Rd. Hampton, GA
2006-2008	Science Teacher Sequoyah Middle School Riverdale, GA

PRESENTATIONS AND PUBLICATIONS:

Koulagna, Y. (2019, April). *Science Teachers' understanding of Formative Assessment and its Practice with Three-Dimensional Teaching and Learning.* Paper presented at the meeting of National Association for Research in Science Teaching (NARST), Baltimore, MD.

Enderle, P., Sengul, O., Koulagna, Y., Grooms, J., & Sampson, V. (Submitted). How implementing and refining an argumentation instructional model can shape science teacher beliefs. *Science Education*

Enderle, P., Sengul, O., Koulagna, Y., Grooms, J., & Sampson, V. (2017, April). *The impact of implementing and refining an argumentation instructional model on science teachers' beliefs.* Paper presented at the meeting of National Association for Research in Science Teaching (NARST), San Antonio, TX.

Koulagna, Y. (2016, June). *My Hair and I: Breaking the Chain and Redefining Beauty*. Presented at the biannual Glorious Hair meeting, Georgia State University, GA

Wade, K., Roman, S., Schoene, M., Koulagna, Y., Spurly, J. (2016, March). *Using Culturally Relevant Pedagogy in Science Classrooms.* Presented at the annual meeting of the National Science Teachers Association, Nashville, TN.

Wade, K., Roman, S., Spurley, J., Schoene, M., & Koulagna, Y. (2016, February). *Using culturally relevant pedagogy in secondary science classrooms.* Paper presented at the Georgia Science Teachers Association conference, Stone Mountain, GA.

Koulagna, Y. (2015, October). Formative Assessment Tasks: *Science Teacher's Perception, Flexibility, and Strategies used in the Process of Creating and Implementing Assessment Tasks.* Presentation at the annual meeting of the Southeastern Association for Science Teacher Education (SASTE), Columbus State University, GA.

Koulagna, Y., Bloxson, K., Hilaski, D. (2013, October). *Teacher Professional Development and Culturally Relevant Pedagogy*. Presentation at the annual meeting of the Georgia Association of Teacher Educators, Jekyll Island, GA.

PROFESSIONAL ORGANIZATIONS

2007 to Present	Georgia Association of Educators
2013 to Present	National Science Teacher Association
2015 to Present	National Association for Research in Science Teaching

SCIENCE TEACHERS' UNDERSTANDING OF FORMATIVE ASSESSMENT AND ITS PRACTICE WITH THREE-DIMENSIONAL TEACHING AND LEARNING

by

YOTAH KOULAGNA

Under the Direction of Dr. Renee Schwartz

ABSTRACT

Formative assessment is increasingly being recognized as a necessary process to improve instruction and enhance learning (Black & Wiliam, 1998; Herman, 2013; Kingston & Nash, 2011). However, the concept of formative assessment is elusive; its definition muddled in policy, practice, and research due to variable goals and perspectives (Bennett, 2011; Black & Wiliam, 1998; Dunn & Mulvenon, 2009). Paired with the dominance of high-stake summative assessment (Furtak & Ruiz-Primo, 2008), formative assessment compared to summative assessment is less appealing to teachers. Compounding to this problem is teachers' lack of formative assessment knowledge and skills especially in the new era of integrating components for three-dimensional teaching and learning advocated by the Framework for K-12 Science Education (NRC, 2012) and the Next Generation Science Standard (NGSS) (NGSS Lead State, 2013). The objective of this qualitative case study was to explore secondary science teachers' understanding of formative assessment and three-dimensional teaching and learning. Three teachers were purposefully selected to participate in the study. A situated lens as the conceptual framework guided the exploration of the research problem and the description of the relationships between specific variables identified in the study. Social constructionism informed the analysis and meaning made from the study, guided the adjustments and decisions taken, and directed the ongoing research to develop a detailed picture of secondary science teachers' understanding and practice of formative assessment and three-dimensional teaching and learning. Data were collected using semi-structured interviews, observations, and documents (lesson plans and assessment tasks). Data analysis occurred iteratively with data collection. Analysis of interview data indicated that teachers understood the concept of formative assessment, and they believed that 3D teaching was a complicated process that required integration of the three dimensions. Analysis of observation and document data indicated that teachers were acclimating to the practice of 3D. They made small changes to their instruction and explored ways to assess students' understanding of 3D learning formatively. They constructed their lessons and assessment task with guidance from the performance expectations of the standard and were mindful of the necessity to integrate the three dimensions. Although cognizant of this synergy, they encountered challenges in the process. Insight from this study has the potential to assist teachers and other stake holders embracing 3D teaching and formative assessment of 3D learning.

INDEX WORDS: Formative Assessment, in-service Science Teachers, Three-Dimensional Teaching and Learning

SCIENCE TEACHERS' UNDERSTANDING OF FORMATIVE ASSESSMENT AND ITS PRACTICE WITH THREE-DIMENSIONAL TEACHING AND LEARNING

by

YOTAH KOULAGNA

A Dissertation

Presented in Partial Fulfillment of Requirements for the

Degree of

Doctor of Philosophy

in

Teaching and Learning

in

Middle and Secondary Education

in

the College of Education

Georgia State University

Atlanta, GA
2019

DEDICATION

This dissertation is dedicated to my two courageous children, Bungha Bengyella and Kapsa Bengyella, you were loving, patient, understanding, supportive to me, and critical to this circuitous journey. You gave me a reason to keep sailing even when the tides were rough.

ACKNOWLEDGMENTS

I will start by thanking the almighty God for giving me good health and a sound mind throughout this journey. To my family and friends, I say thank you for giving me unconditional support and all the food and encouragement, for listening when I needed someone to talk to, and for reassuring me whenever I stumbled.

This academic pilgrimage has been met with turns and twists and I am humbled and thankful to everyone who has counseled me in any capacity at each level of achievement, especially my dissertation committee. Your insightful feedback is very much appreciated. To Dr. Gertrude Tinker Sachs, you extended a hand and lifted me up when I was down, and you continuously coached and cheered me on. You provided a strong shoulder for me to lean on and edge forward, for that I say than you. To Dr. Renee Schwartz, my committee chair, you boarded the train when I was experiencing academic disorder and you gave me direction towards research and scholarship. Thank you for your patience and for guiding me towards the finish line. To Dr. Patrick Enderle, your office was always open, and you were always ready to provide me with expert advice and the motivation I needed; thank you immensely. To Dr. Jodie Kaufmann, you said “yes I will help you” with no reservation and your advice was critical and timely to the last leg of this expedition. To my dream team, again I say *mercy beaucoup*.

PLEASE NOTE: An editor has not been used in the construction of this dissertation.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

1. PROBLEM

Background

In the past, most teacher education programs did not formally offer assessment as a course of study (Darling-Hammond & Bransford, 2005) and most pre-service and in-service teachers learned about assessment from their days as students and from seeing others teach or according to Lortie (1978), from their apprenticeship of observation. In addition, the yearly in-service trainings and the sporadic professional development do little to fill the void for most teachers. Recent research also indicates that about 60% of pre-service teachers receive little to no training on formative assessment and only 40% receive minimal training (Stevens, 2012). Most teachers therefore enter the classroom lacking the knowledge and skills for practicing assessment. This deficit exists as teachers have become experts in summative assessment due to pressure from high stake testing (Darling-Hammond & Bransford, 2005).

Informing teaching and learning

High achieving countries focus more on formative assessment to produce noticeable outcomes for school-based assessment tasks (Darling-Hammond & McCloskey, 2008). Formative assessment was first authored in the literature by Bloom in 1969. He defined formative assessment as an evaluative process necessary "to provide feedback and correctives at each stage in the teaching-learning process" (p. 48). Twenty years later, Black and Wiliam (1998), described formative assessment as encompassing "all activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (pp. 7-8). Formative assessment is thus a descriptive, interpretive, and steering process which teachers and students engage in, to elicit and act upon evidence about the teaching and learning. It employs feedback as correctives/adjustments at every step of the teaching and learning journey (Black & Wiliam, 1998; Bloom, 1969) and uses evidence to make better founded decisions about teaching and learning than decisions without such evidence (Black & Wiliam, 2009). Formative assessment can be informal or formal. During informal assessment, the teacher acts impromptu or on the fly (Ruiz-Primo & Furtak, 2007) to students' ideas and thus provides evidence for continuous learning. Assessment is formal when evidence is gathered to plan for lesson and used before and/or after instruction and learning, therefore provides evidence of students' learning. Emphasis on classroom assessment encourages the need to clarify in advance the lesson's objective for students (Where they are going), to assess and inform students on their present situation (Where they are), and to provide students with information as feedback or road map of how to meet their objective (How to get there) (Black & Wiliam, 1998; Sadler, 1989). Teachers in the process gain revealing information about students' current and evolving progress to help them adjust instruction and improve students understanding of concepts (Yin et al., 2008).

Increasing acknowledgement and surmounting evidence exist today on the role formative assessment plays in improving instruction and helping students achieve concept mastery (Herman, 2013). Formative assessment elicits evidence of learning at each point of instruction to understand "how student is evolving as a learner" and "how to assist the learner to his/her pathway to mastery" (Sadler, 1989, p. 121). Formative assessment is timely as the teacher uses the information immediately to make instructional adjustments (remediate, reteach concept and skills) and help build local capacity to sustain this shift overtime (Black & Wiliam, 1998; Sadler, 1989). This assessment provides varied opportunities for students to apply the knowledge and skills learned to solve problems and raise expectations (Black & Wiliam, 1998; Darling-Hammond & McCloskey, 2008; Kang, Thompson, & Windschitl, 2014; McClellan, 2004). The low-stake nature of FA couple with its feedback potential help students develop confidence in the ability to express their understanding freely, it gives voice to all students as they engage in inquiry, and it welcomes and value students' experiences in discussions and development of knowledge (Ruiz-Primo & Furtak, 2007). Formative assessment is thus a fair assessment as it levels the playing field for all students. Formative assessment projects forward as a safeguard and is hence an asset to teaching and learning. Formative assessment can improve two-way communication that is reflexive and relational between teacher and student giving them shared authority over learning (Buck, Trauth-Nare & Kaftan, 2010; Stiggins, 2002). It is a partnership between teacher and student, each equally responsible for the richness of assessment conversations. It provides teachers and students with opportunities to identify areas of weaknesses to minimize them and increase areas of strengths to master curriculum and improve performance (Chappius & Chappius, 2008).

Feedback potential

Formative assessment is grounded on the theoretical model of learning and its regulation, using interactive descriptive feedback as a vehicle (Black & William, 2009; McTighe & O'Connor, 2009). Formative feedback transforms learning by helping students internalize features of good work. It provides students with a clear vision of targeted skills, appraises current progress, and explains how to improve (Rushton, 2005; Shepard, 2005). Shepard suggests that reflexive constructed feedback guides students' judgement, affects self-perceptions of competence, focuses on developing habits of thinking for deep learning not rote memorization that comes with surface learning. Feedback fosters cooperative learning, active engagement amongst peers and development of self and peer assessment skills. Formative feedback should be timely, non-evaluative, supportive, and specific or relevant to future assignments. Feedback can also be delayed as with written responses. It can involve verification of accuracy, explanation of correct answer, or serve as hints and prompts (Shute, 2008). The power of feedback (Hattie & Timperley, 2007) is thus in its directives to eliminate the gap between student's actual and reference levels of conceptual understanding. However, pressure from grading and teacher workload often result in a kind of feedback that suffers from timeliness, quality, and quantity, losing its effectiveness as a formative tool (Glover & Brown, 2006). Feedback renders previous formative assessment obsolete since students use the information to adjust their thinking and progress, therefore, teacher must reassess to continuously adjust and keep teaching and learning on track (Heritage,2011). This feedback loop is thus necessary to provide equal opportunity to meet the intellectual and social needs of all students (Sadler, 1989).

During feedback conversations, the teacher facilitates dialectic discourse that shapes participation which in turn is shaped by the discourse (Anderson, Zuiker, Taasoobshirazi, & Hickey, 2007). The teacher elicits and recognizes students' ideas and communication skills and uses the evidence to adjust instruction, while students participate in assessment conversations as critiques to their peers and to their own ideas. Formative assessment thus is concerned with appraising the quality of responses that can shape and improve students' competency therefore eliminating the inefficiency of guess-work learning (Sadler, 1989). Formative assessment according to Heritage (2011), is "an approach to teaching and learning that uses feedback as centerpiece in a supportive classroom context" p.19) to move the teaching and learning forward.

Three-dimensional teaching

Much is expected of what U.S. science teachers should know and can do based on popular discourse on students' achievement in science, engineering, and technology (National Research Council- NRC, 2010). Fortunately, teachers' expected facility with different knowledge, how students learn concepts, and how to teach, is aided by continuous scholarship on teaching and learning science. The National Research Council Framework's vision is for students to acquire knowledge and skill in a sequence of stages that develop their understanding of aspects of three-dimensional practices in each standard (NRC, 2012). The vision requires that each standard and its performance expectations combine three dimensions:

1. Science and Engineering Practices (SEP)- teachers should explain the skills and knowledge scientists engaged in to investigate the natural world such as building models and theories, and the skills and knowledge that engineers use to design and build models and systems.
2. Crosscutting Concepts (CCC)- teachers should make explicit the concepts that provide organizational schema to link knowledge from the different domains of science. This include cause and effect; patterns, similarity, and diversity; scale, proportion, and quantity; energy and matter; structure and function; systems and system models; stability and change.
3. Disciplinary Core Ideas (DCI) - teacher's instruction and assessment should be grounded in the core ideas of the domain. These ideas must have a broad importance or can organize concepts; be an instrument for understanding or investigating complex ideas and solving problems; consider students interest and experiences or individual and societal needs; be teachable and learnable across grade level and with increasing sophistication (NRC, 2012).

Involvement in scientific discussions deepen students' insight, curiosity, and responsibility to figure out why certain aspects of the natural world works. The rationale for integrating the three dimensions is that to fully understand science and engineering ideas, students must engage in inquiry and discourse practices to develop and refine such ideas; students need a specific context to develop competence in practices; students use crosscutting concepts to make connections between and as intellectual tools across discipline (NRC, 2012).

Science learning is three-dimensional. To solidify students' learning, repeated opportunities to participate in scientific thinking and practices must be provided to enable students to slowly build an understanding of how new knowledge assimilates with old experiences (National Academic of Science, Engineering, and Medicine [NASEM], 2017). Robust learning mandates coherence among instruction, curriculum standards, and assessment. Aligning the teaching, curriculum goal, and assessments to past and future grades help expand students' understanding of CCCs and DCIs, and their skills with SEPs (NASEM, 2017). This synergy helps traces the path through which students make progress by explaining what they are expected to know and be able to do. Reform documents suggest that teachers anchor their instruction in phenomena so students can actively engage in science thinking to be able to make connections and understand how and why science ideas are important (NASEM, 2017; NGSS, 2016). Engaging all students in three-dimensional teaching requires that teachers strive to structure instruction that, is student centered, weaves the three dimensions, is flexible for student to explore, is cumulative, provides repeated opportunity to engage students with ideas, uses engaging phenomena that motivate students to explain it, and caters to the needs of all students (NASEM, 2017).

The formative notion of looking forward and continuously assessing where students are relative to the standard or learning goals is possible with the teacher first mapping the learning sequence (Herman, 2013). Even though Georgia did not adopt the NGSS recommendations, Georgia Standards of Excellence are heavily reliant on NGSS. The Next Generation Science

Standards and the Framework for K12 science education considered learning as a trajectory through which students evolve throughout the unit, year, or K12 while participating in three-dimensional learning (NRC, 2012). The cognitive model of how science learning progresses is based on sequencing learning to provide a baseline for diagnosing and evaluating the gap (Heredia & Furtak, 2014; Heritage, 2008). Learning progression according to Shepard (2019), is a detailed model of learning composed of the learning goal, an intermediate step, and the instructional strategy for achieving the goal. To be effective, learning progression requires a simultaneous design of the instructional activities, the assessment, and the teacher's learning support. This pathway to mastery according to NGSS is a key theme in science learning and coherence in science education (NRC, 2012). Sequencing provides information about how student's understanding and ability to apply scientific concepts develop more sophisticated overtime and helps teacher know where students are at each stage of learning with respect to success criteria. Formative assessment practice must be linked to the learning goal through learning progression to support classroom assessment for three dimensions (Shepard, 2019).

Teacher thinking and Practice

A shift from external/summative forms of assessment to an increasing emphasis on classroom assessment is not a new phenomenon (Black & Wiliam, 1998; NRC, 2001). However, formative assessment practice is scarce in science classrooms. It is understood that teachers' beliefs system is influenced by their beliefs in science, in self, and in teaching. Beliefs and attitudes also shape the way teachers interpret and respond to change and challenges (Jones & Leagon, 2014). Belief about self (self-efficacy) according to Jones and Leagon (2014), influences teachers' behavior towards practice/instruction, motivation, success of professional development, and towards educational reform. They are confident that teachers' prior experiences including success and failure histories, and feedback from others influence self-efficacy or the ability to get things done. Science teachers will need more than belief in ability to embrace the instructional principle of three-dimensional teaching and be willing to monitor the teaching and student learning minute by minute.

Throughout their lives as learners, teachers have established and nurtured their perception and beliefs about teaching. They have in the process, personally constructed their own knowledge and beliefs about teaching that influence “the structure of the classroom, the way the curriculum is interpreted, and how instructional practices are enacted” (Jones & Leagen, 2014, p. 832). To change their disposition towards reform messages, science teachers must be self-motivated, believe in the potential of the new strategy, and be provided with necessary resources (Borko, Davinroy, Bliem, & Cumbo, 2000). It takes a long time for teachers to build a practice that works and changing to a new one is very difficult. Asking a teacher to embrace a new way of doing things, according to Lee and Wiliam (2005), “is like asking a golfer to change his or her swing during the tournament” (p. 13). They believe that change is a slow process because teachers want to keep what has been successful. Teachers need continuous support to integrate new strategies into their existing practices. Therefore, before high school science teachers can embark on this reform journey, credible evidence that formative assessment does improve practice and students’ learning must be presented (Lee & William, 2005) and teachers must develop an understanding of the nature and process of formative assessment (Dunn & Mulvenon, 2009) with three-dimensional learning. If teachers lack the knowledge and support, and assessment tasks are simply inserted as special activities into their lesson especially in the new context of three-dimensional teaching, they will have difficulty enacting them. Given that teachers’ personalities as well as their classroom settings has the power to shape and constrain practices, the pattern may become automatic resulting in resistant to reflection or change

(Putnam & Borko, 2000). To help teachers overcome the reluctance of implementing formative assessment of three-dimensional teaching and learning, we need to understand teachers' status of such practices couple with the willingness to change their view of pedagogy and the curriculum they are responsible for implementing (Rushton, 2005).

Problem statement

The concept of formative assessment is elusive, its definition muddled in policy, practice, and research due to variable goals and perspectives (Bennett, 2011; Black & Wiliam, 1998; Dunn & Mulvenon, 2009). Teachers need a clear understanding of formative assessment, sound formative assessment strategies and tools to obtain valid inferences of students' learning. While teachers tacitly assess students' understanding daily, assessment is usually reserved for the end of the lesson or unit and feedback information are rarely used to adjust instruction (Furtak & Ruiz-Primo, 2008). Assessment as such is viewed mostly as a tool to judge what students have been taught (Darling-Hammond & Bransford, 2005), by concentrating on the limited properties of test with little connection to the learning experiences of students (Black & Wiliam, 1998). Although a shift from summative towards the more impactful formative assessment has been endorsed (Black & Wiliam, 1998), its practice is still sparse.

Evidence supporting the positive impact of formative assessment on student learning (Black & Wiliam, 1998; 2009) has come mostly from laboratory studies or unreliable records with no ecological validity (Yin et al., 2008) thus not generalizable to real-life science classroom setting. Likewise, empirical studies have focused predominantly on students' thinking, learning, and nature of knowledge gained, but little related to teachers understanding and how they create learning experiences aligned with formative assessment (Putnam & Borko, 2000). Among the few studies that investigated science teachers' formative assessment practices, some created tasks for teachers to embed into their lessons (Tomanek, Talanquer, & Novodvorsky, 2007; Wiliam, Lee, Harrison, & Black, 2004; Yin et al., 2008), while others gave teachers one-time limited training to create structured tasks to embed into their lesson (Kang et al., 2014; Metin, 2013). Though these are positive steps towards highlighting the importance of tasks in formative assessment, little effort has been made to uncover high school science teachers' understanding of formative assessment especially as it relates to assessing three-dimensional teaching and learning.

The Next Generation Science Standards is changing the way teachers think about teaching and learning. This change requires a shift from assessing knowledge of content to assessing understanding of core ideas using crosscutting concepts and appropriate science and engineering practices. It is obvious that teachers could benefit from developing new strategies as they transform standards into rigorous science tasks that require students to perform beyond basic competency level by focusing on explanatory reasoning and exploratory abilities with higher order thinking skills (NRC, 2001). The payoff will come when teachers can link the curriculum standard to their instruction and assessment (such that the goal of the lesson and the curriculum align to past and future grades) and weave the dimensions together seamlessly for the outcome to be cumulative (NASEM, 2017). Teachers must think differently about assessment for three-dimensional teaching and learning. The new formative assessment will gage how students use SEP in the context of CCC and DCI, and use a variety of challenging tasks (that provide specific and varied evidence of students' status), so the teacher can adjust instruction to focus on students' progress along the learning pathway. Tracking the path towards mastery by explaining what students are expected to know and be able to do, based on learning goals (for each topic and level) would help the student learn to transfer knowledge and reasoning to gain information from one domain to another (NASEM, 2017). However, information on three-dimensional learning is theoretical with minimal verifiable development available to guide teachers on how to proceed (Fick, 2017). Furthermore, the NRC Framework's recommendation to create new system of assessment to monitor student's understanding and progress of three-dimensional learning has yet to yield evidence to guide teachers (Fick, 2017; Harris et al., 2015; Herman, 2013). Although assessments from testing programs assess the practices and challenging concepts, their designs do not factor three-dimensional learning yet (NASEM, 2017). Therefore, teachers must rely on their own experiences of how learning evolves to develop assessment that integrate SEP, CCC, and DCI to facilitate three-dimensional teaching and learning. To encourage classroom formative activity, information on science teachers' knowledge and pedagogy of three-dimensional teaching is essential.

Purpose

The purpose of this study was to explore science teachers' understanding of formative assessments and how they elicit and interpret students' integrated science knowledge to adjust instruction and improve students' three-dimensional learning. The study also explored teachers understanding of three-dimensional teaching and the strategies used and/or challenges used to implement formative assessment. Another reason for the study is that in order to help teachers integrate FA into their practices, it must be known what they understand and do. Therefore, this study identified formative assessment strategies used by these teachers to promote students' understanding and application of science concepts and to understand inherent obstacles, challenges, and necessity for implementing formative assessment with three-dimensional teaching.

Research Questions

The following four questions guided the direction of this study:

1. What are high school science teachers' understanding of formative assessment?
2. How do high school science teachers understand three-dimensional science instruction?
3. How do high school science teachers practice three-dimensional teaching?
4. How do high school science teachers practice formative assessment of three-dimensional learning?

Significance of Study

The previous sections provided some justification that science teachers lack the knowledge of creating formative assessment tasks and struggle with its implementation (Darling-Hammond & Adamson, 2013; Metin, 2013). Besides, there are no verifiable formative assessment strategies for monitoring and improving three-dimensional teaching and learning (Fick, 2017; Harris, 2015; Herman, 2013). Because teaching, learning, and assessment are social practices involving the construction of meaning, formative assessment is vital in tracking how and when learning is happening with three-dimensional teaching. This study is significant because it provided an image of secondary science teachers' perspectives and practice of formative assessment in the era of three-dimensional teaching and learning. Gaining an understanding of science teachers' knowledge and practice of formative assessment is necessary. The information could provide valuable insight towards packaging and tailoring the appropriate grain size for strategies and resources that could lead to significant changes in teaching practices and student learning (NRC, 2012; Wiliam et al., 2004). Administrators, professional development facilitators, and teacher education programs could also gain practical information to support teachers with three-dimensional teaching and assessing learning. Furthermore, insight from exploring these questions could provide clarity into the concept of formative assessment, as minute learning goals of summative assessment, and reposition teachers' views of formative assessment as an additional burden (Jenkins, 2004). Individual research collectively endeavors to build on the incremental science knowledge base. As part of a collective effort, this study also aimed to contribute to laying the groundwork for future research and development of formative assessment for three-dimensional teaching and learning.

Assumption and Limitations

When teachers track students' past, current, and evolving progress to provide students with corrective feedback and adjust their instruction, teachers can improve their practice as well as students learning. Reform based strategies like quality formative assessment could provide teachers with opportunities to practice developing scientific explanations for phenomena and gathering critical information about student's learning to continuously guide their instruction. Teachers intuitively elicit students' ideas during instruction and as such are not purposefully using the feedback information to clarify the objective of the lesson, assessing students' pre- and present conceptions, and providing a road map of how to meet their objectives. This study was limited to high school science teachers' knowledge, experience, and flexibility, to implement and formatively assess three-dimensional practices.

Overview of the Study

This chapter provided background information and a rational for conducting the study. Chapter two explored the literature on formative assessment in science classrooms to provide a detail account of the status of formative assessment, what issues has been investigated, and what still need further exploration. Chapter three described the methodology that guided the direction of this study and the design that lead the quest and collection of necessary data to answer the research questions. Chapter four described and developed cases for the three teachers and created a cross case synthesis of the cases for commonalities and differences that emerged. Chapter five answered the research questions and discussed the findings and suggestions for future studies.

2. LITERATURE REVIEW

Introduction

According to the National Research Council's report, it is possible to close and even eliminate the achievement gap in US K12 science education (NRC, 2012). One way to improve all students learning documented by many is through formative assessment (Anderson & Palm, 2017; Black & Wiliam, 1998; Herman, 2013; Kingston & Nash, 2011). Despite this growing recognition for formative assessment, teachers are reluctant to practice it due to- 1) dominance of high-stake summative assessment (Furtak & Ruiz-Primo, 2008); 2) teachers lack the knowledge of creating assessment tasks and struggle with its implementation (Anderson & Palm, 2017; Darling-Hammond & Adamson, 2013; Heredia et al., 2016). Heritage (2008) argued that learning is a developmental process. Understanding how students may progress throughout a domain can help teachers develop formative assessment abilities (Bennett, 2011; Heritage, 2009; Herman, 2013). Therefore, teachers need knowledge of the pathway along which students evolve across the unit and where they are concerning the learning goals as they participate in practices to develop core ideas in science (NRC, 2012). However, there are no verifiable formative assessment strategies for improving teachers' three-dimensional teaching and learning (Fick, 2017; Herman, 2013). Insight into how science teachers collect, analyze, and interpret data of students' evolving sophistication of disciplinary core ideas is essential. The goal of this review was to critically explore and synthesize current information on formative assessment and three-dimensional teaching and learning, to identify what is present in the literature and what needs further exploration. The analysis of literature was divided into two parts; the first section includes criteria for including and excluding articles, search engines, terminologies, and a table to provide a summary visual of the articles reviewed. The second section comprises the body of the paper containing the literature discussed under the following headings- 1) Definition of formative assessment; 2) Embedded formative assessment task 3) Classroom conversations and response trajectory 4) Perception on and nature of feedback loop 5) Professional development and teacher change 6) Formative assessment in relation to teacher knowledge 7) Sequencing learning for formative assessment 8) Three-dimensional teaching and learning. This review concluded with significant findings, limitations of the studies and the gap, implications, and thoughts for the future, and suggested questions for further investigation.

Criteria for inclusion and exclusion

The articles included in this review met the following criteria

- Time frame was after Black and Wiliam's (1998) seminal work (given that relevant sources were included in their review)
- Sources were primary studies reported by the original researchers
- Population of science teachers and their students and exception given to teachers and students of mathematics and other discipline if the study provides unique insight to FA.
- The level spans from elementary to higher education, to expand the context and sources
- Data type was either self-collected data or data as part of a team of a project
- Language of study was English
- Studies that reported findings on at least one feature of FA such as descriptive feedback, self-assessment, scaffolding, learning progression, goal and success criteria, collaboration or classroom interaction and dialogue.
- Studies on professional development and others that enhance teachers' practice and quality of FA, and studies that examine teachers' thinking and practice of FA

- Studies relevant to FA, provided relevant insights, added to the academic discourse, and provided implication and recommendation for future research.

The Following Criteria Were Used to Exclude Articles

- Studies that discuss FA in general in other disciplines
- Studies that mainly compares FA to summative assessment with no input or output variables
- Literature reviews on FA although their reference lists served as sources for potential articles to be included

Search engine and terminology

Terminologies identified from these review articles included formative assessment, classroom assessment, assessment for learning, three-dimensional teaching/learning, and secondary science. These terms were used to search for articles in EBSCO Host-ERIC, Academic Search Complete, Education Full Text, and Google Scholar. This review started with a Google scholar search for literature reviews on FA. The result yielded three original articles whose reference lists served as a starting point for relevant terminologies and article sources. A challenge with the article search using these key terms was that they are not used uniformly to address the interest of this paper. FA will be used as an umbrella term to include an assessment where the primary purpose in its design and enactment is to improve instruction and foster student learning (Black, Harrison, Lee, & Marshall, 2004).

[Table 14.- see Appendix]

Definitions of Formative Assessment

The use of formative assessment remains an mystery as many different definitions permeate the education literature (Black & Wiliam, 1998; Shepard, 2005). There will continue to be a persistent shortage of scientific evidence of its impact so long as vagueness in the constitutive and operational definition exist (Dunn & Mulvenon, 2009). Researchers use different terminologies to mean the same thing, and sometimes, the same vocabulary to say different things (Black & Wiliam, 1998; Meyer, 1992). A clear definition is essential to help researchers' document effectiveness and compare its impact across studies and transfer the knowledge to other contexts. Early on, Bloom (1969) defined formative assessment as an evaluative process "… to provide feedback and correctives at each stage in the teaching-learning process" (p. 48). Chappuis and Stiggins (2002) defined a formative assessment as an assessment designed to monitor students' progress during the learning process. While Black and Wiliam (1998), described formative assessment as encompassing "all activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (pp. 7-8). Assessment in their definitions was interpreted as instruments for collecting information about students' progress during the learning process (Dunn & Mulvenon, 2009) and adjust instruction during instruction. A more inclusive definition by Chappuis et al. (2012) goes beyond looking at what happens during instruction to include teachers and students assessing students' strengths and weaknesses, knowledge, and skills before, during, and after instruction. The Framework for K12 education's goal provides a similar definition for assessment as, "an ongoing activity, one that relies on multiple strategies and sources for collecting information that bears on the quality of student work and that then can be used to help both the students and the teacher think more pointedly about how the quality might be improved" (NRC, 2001, p. 30). The teacher and students recognize and respond to the student's learning, and the teacher adjusts instruction to enhance learning and teaching during learning and teaching (Bell & Cowie, 2001). The trend that keeps repeating in the literature is that "it is *not the instrument* that is formative; it is the *use of the information* gathered, by whatever means, *to adjust teaching and learning*, that merits the "formative" label" (Chappuis, 2009, p. 4). Black and Wiliam (2009) argue that, a

> practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded than the decisions they would have taken in the absence of the evidence that was elicited (p. 7).

Embedded Formative Assessment Tasks

Teachers have become experts in implementing standardized tests over the years, but the same does not apply to formative tasks. Model for teachers' reasoning about assessment requires a combination of cognitive, observational, and interpretive skills (NRC, 2001). The essence of creating assessment tasks is for teachers to explore different factors or reconstruct formative assessment strategies to accommodate their classroom style, to inform them about students' evolving understanding continuously. The teacher assesses the gap and provides timely feedback to students on how to close the gap in their knowledge (Black & Wiliam, 2004a). These mini-assessments are coordinated with end-of-unit assessments to signal a unit's organizational goal and give direction to the teacher (Shavelson et al., 2008). Thus, in creating new assessment tasks, the skills and knowledge incorporated should be that which is assessed (Black & Wiliam, 1998; Tomanek et al., 2007). Formative assessments tasks as fragments of formal assessments are embedded at strategic junctures in the curriculum or unit to create goals or sub-goals directed towards teachable moments before the student progresses to the next lesson (Yin et al., 2008).

The studies reviewed in this section highlight the complexity of constructing and challenges in implementing embedded formative assessments.

Wiliam et al. (2004) conducted a study to examine the achievement of secondary school students in classrooms where teachers made time to develop formative assessment strategies. In this study, 24 mathematics and science teachers were selected, trained with formative assessment strategies/techniques, and given a choice to choose which strategies to use. They assessed students with assessment instruments. The authors acknowledged that the test lacked curricular validity (did not measure necessary concepts nor were aligned with the curriculum). Observations and results of the interquartile range in effect size on test results revealed that teachers' practices were slow to change, and any observable change was towards the end of the year. Wiliam et al. (2004) attributed the quality of formative assessment to the teachers' expertise level and suggested that using embedded formative assessment improved students' achievement in externally mandated assessments. However, Wiliam et al. (2004) cautioned the acceptance of their results based on different units of comparison. They suggested further research on students' achievement in classrooms where teachers formally embed assessments into their lessons.

Yin et al. (2008) conducted a similar study to that of Wiliam et al. (2004) but embedded a formative assessment aligned with the curricular goal and summative assessment with a different outcome. Six experimental teachers taught a formative assessment embedded curriculum. The purpose of this quantitative exploratory study was to determine whether the embedded formative assessment task can improve students' motivational beliefs, students' achievement, and conceptual change. The results of a motivational belief questionnaire and achievement assessment indicated that embedding formative assessment in the curriculum had no significant influence on students' motivation, achievement, and conceptual change. Yin et al. (2008) noted that trained teachers varied in degree of implementation of embedded formative assessment which, affected students' outcomes differently. Their results revealed the difficulty of implementing formative assessment even after providing teachers with training on how to use the information to improve teaching and student learning. They cautioned about solely embedding formative assessment into the curriculum without inviting teachers to participate in its design and provide follow up in-progress training.

In a similar study, Tomanek et al. (2007) explored science teachers' reasoning associated with task selection or evaluation of factors used as a planned formative assessment. The goal was to identify teachers' knowledge and beliefs that ground their assessment decisions. Data were collected from 24 first and 27 second-year teachers in a Science Teacher Preparatory Program (STPP) and 41 experienced science teachers who usually collaborate with the STPP faculty as mentors. They used formative assessment probes as a data collection instrument. However, the reliability and validity of their instrument was questionable and used as a pilot tool. Descriptive analysis of probes for teachers thinking revealed that task selection and evaluation was based on two themes: (1) "Characteristics of the tasks" and (2) "characteristics of students or the curriculum" (Tomanek et al., 2007, p. 1119). They found no relation between being able to judge the level of thinking demanded of the task and teachers' training or experience. However, selecting tasks based on students or curriculum characteristics was related to teaching experience. They implied that (1) teachers reasoned with factors that sometimes work against the selection of tasks with the potential of assessing students' understanding of concepts. (2) Prospective and experience teachers must be afforded the opportunity to question their beliefs about assessment and interpretation of evidence of student understanding.

These studies highlight teachers' reasoning about selecting assessment tasks and the challenges in creating and implementing embedded formative assessment. Wiliam et al. (2004) and Yin et al. (2008) findings on the relationship between teacher enactment and student achievement indicated that teachers had variable expertise and affected students' success differently. While Wiliam et al. (2004) reported a significant improvement in students' achievement with embedded assessment tasks, Yin et al. reported the contrary. Both studies found that experience alone is not enough for the effective implementation of formative assessment. Tomanek et al. (2007) reported that training or experience had no relationship with being able to judge the level of thinking of tasks. These studies suggest that teachers need more training in constructing or participation in co-constructing and enacting formative assessment tasks.

Complexity of Classroom Conversations and Response Trajectory

Evidence in learning science indicates that affinity is more towards talking science than reading science (Lemke, 1990; Roth, 2005). The dynamics of talk provides an interactive medium for exploring and knowing about the world scientifically. Therefore, whole-class discussions create opportunities for students and peers to question and provide a rationale for scientific claims and for the teacher to solicit, monitor, and enhance students learning. Classroom talk is vital to many approaches to learning but difficult to coordinate because of the challenges to manage rich classroom discourse and inquiry activity in unison (Anderson et al., 2007). These daily assessment conversations or instructional dialogues were embedded in an activity currently taking place in the classroom. It can allow the teacher to gather information about the status of students’ conception, language use or communication skills, mental models, or used as strategies to guide instruction (Ruiz-Primo & Furtak, 2006). However, classroom talk does not reveal students' complete conceptions to the whole class. Adding another component of curriculum-embedded assessment as written explanation (silent talk) could provide a broader definition of feedback that addresses each student's need (Furtak & Ruiz-Primo, 2008). Thus, formative assessment as a tool can be employed at any level of student-student-teacher interaction during daily classroom talk to improve students' conceptual understanding and assist teachers continuously gain insight about students' level of understanding (Bell & Cowie, 2001; Furtak & Ruiz-Primo, 2008). This section described studies that demonstrate the impact of formal (planned) and informal (unplanned) formative assessment in the classroom using various kinds of discourses.

Focusing on discussions between students, Anderson et al. (2007) conducted a study to understand how research efforts to document discursive classroom routines informed the process of advancing participation in scientific inquiry and performance on high-stake achievement. Using a situated lens, Anderson et al. (2007) explored "individual student performance on individual test items as examples of specific types of discourse, allowing a coherent examination of transfer of understanding across very different ways of knowing" (p. 1742). Videotape discussions from groups of 11^{th} and 12^{th} grade students and their pre-and posttest results were analyzed. The results highlighted students' engagement in classroom conversation, the role of answer rubrics, and the teacher facilitation that was better coordinated to scaffold more productive discursive trajectory classroom talk in the second year. Anderson et al. (2007) noted an improvement in teacher providing informal feedback during classroom conversations and more productive discussions with the use of answer rubrics. Also, a varying degree of learning was reflected in group discourse on quiz feedback and on gains in the examination and test. The implication from this study is that the role and nature of group discourse couple with teacher intervention affect the quality of students' movement along the discursive trajectory from informal formative feedback conversation to more formal assessment activities.

Looking at discourses between students and teachers, Ruiz-Primo and Furtak (2007) explored how students develop an understanding of concepts during daily whole-class conversations. They used the Elicit, Student respond, teacher Recognize and Use (ESRU) model to distinguish the quality of informal assessment practices across teachers and to determine whether this quality can be linked to student performance. Videotapes of classroom conversations were collected from three middle school teachers trained in the implementation of the FAST (Foundational Approach for Science Teaching) curriculum. The analysis of discourse transcripts indicated a range of informal assessment frequencies, from incomplete (ERS or IRE/F) to complete conversation cycles (ESRU). They made the inference that better informal assessment practices could be linked to better student performance, and the ESRU model was useful in capturing differences in teachers' informal assessment practices. However, this conclusion is drawn from data for only a single teacher.

Furtak and Ruiz-Primo (2008), in a similar study, added a written portion to students' responses to classroom talk to capture students' complete conception of knowledge status, and compare this status with the learning goal. Their study examined the relative utility of the formal and informal functions of four types of formative assessment prompts in eliciting middle school students' ideas about sinking and floating through written responses and classroom discussions. Videotape of written assessment and discussion implementation were collected from four Romance project teachers. Data analysis indicated that all prompts elicited a high percentage of students' ideas at the expected levels, but below-level conceptions were expressed more in writing as compared to discussions. Many more students provided responses containing multiple conceptions in writing as compared to students' responses in whole-class discussions, and some prompts were more successful than others in eliciting a range of students' ideas. Furtak and Ruiz-Primo (2008) noticed that teachers were not using whole-class settings efficiently to elicit students' conceptions and suggested that teachers need more familiarity with prompts for effective implementation. They recommended that teachers should base their judgment on students' conception of both written work and classroom discussion. Besides, future research should explore in greater depth ways teachers implement different assessment prompts and the extent of feedback provided to students about the conceptions elicited in the prompts.

Ruiz-Primo and Furtak (2006) compared how four middle school science teachers from the same experimental group used questioning as an informal formative assessment method to measure students' learning. The purpose was to understand how informal assessment looks like in the context of scientific inquiry, the different levels of informal assessment practices, and whether these different levels were linked to levels of student learning. Data were collected from videotapes of classroom conversations and responses from pre-posttest assessments, and embedded assessments prompt. The analysis results indicated that 1) teachers used complete cycles only 26% of the time and most cycles were 95% epistemic in nature 2) the pattern of change in students' post-test results reflected a change in teacher's informal formative assessment practice 3) teachers using complete ESRU cycle had students with high performance. Ruiz-Primo and Furtak (2006) conclude that it is important to provide teachers with tools with which to respond in immediate and effective ways. They suggested that future studies were needed to revise scientific inquiry domains but did not indicate challenges/limitations with their study. Also, future studies could use ESRU coding techniques to explore discourse in the context of embedded assessment. Ruiz-Primo and Furtak (2006) asserted that evidence from their study could be used to design assessment courses for preservice teachers and professional development on informal formative assessment strategies.

Hickey et al. (2012) study intended to promote meaningful participation in the discursive construction of shared domain knowledge and improved achievement with a design strategy made up of three different levels of assessment (informal/close, semi-formal/proximal, and formal/distal level). Through a situated lens, Hickey et al. (2012) rationalized that using different forms of assessment makes using and aligning them easy, such that an assessment can serve a formative function in one situation and summative function in another. Data were collected through informal observations and videotapes of feedback conversations then aligned to establish collective accountability in students' participation. A sociocultural lens was used to analyze these data for disciplinary discourse and interaction for better understanding. Results indicated that enhancement in feedback conversations parallel gains in proximal exams and distal test. Hickey et al. (2012) claimed that their study provided support for an assessment design model that embeds informal formative assessment into inquiry-oriented activities, assesses and improves the activity-assessment combo, and evaluates the combo against externally developed standardized test items. Limitation of their study was in the challenge of using researcher-developed assessment to evaluate formative assessment practices and failure to clarify the difference between formative and summative functions. Hickey et al. (2012) suggested that the introduction of individually oriented formative assessment into existing classroom instruction may lead to reduce learning.

These studies revealed mixed results with teacher intervention. Increase facilitation improves teacher's domain knowledge, scientific inquiry, and achievement (Anderson et al., 2007; Hickey et al., 2012). Frequent use of incomplete cycle conversations was observed more than complete cycles. However, increasing the use of complete cycle conversations leads to improved teacher practice and students' performance (Ruiz-Primo & Furtak, 2006, 2007). Different assessment prompts elicited different conceptual ideas from students (Furtak & Ruiz-Primo, 2008). The studies in this section focused on teachers' practice with different forms of assessments, and the challenges observed may stem from a lack of teachers' understanding of formative assessment.

Perception on and Nature of Feedback

Formative assessment provides fuel for teaching and learning in the form of descriptive feedback along the way (McTighe & O'Connor, 2009). Feedback is critical for the formative evaluation as it provides students with a clear vision of the skills to be learned, appraises current progress, and explains how to improve (Rushton, 2005; Shepard, 2005). Feedback thus addresses the goal of the lesson by aiming to answer three questions. Where is the student going? Where is the student now? And "how to get there"? (Hattie & Timperley, 2007). Reflexive constructed feedback guides students' judgment and affects self-perceptions of competence and focuses more on assessing deep learning and away from rote memorization that comes with surface learning. It fosters cooperative learning, active engagement amongst peers, and development of self and peer assessment skills (Sadler, 1989). Feedback, in general, gives information about how successful or not something has been done. Thus, the power of feedback arises from its ability to provide that factor necessary to eliminate the disparity between student's present and anticipated level of conceptual understanding (Hattie & Timperley, 2007). Meeting students' physical, intellectual, and social needs require practice in a supportive environment involving a feedback loop (Sadler, 1989). Feedback is a system-control function that connects both the need of the teacher and the student for better instruction and deeper learning. The teacher checks for readiness, diagnoses, and remediates, while the student monitors the strengths and weaknesses of their performance to recognize and reinforce aspects associated with high-quality work. The formative assessment generates feedback promptly to give students the opportunity to self-regulate and can take different forms depending on the context (Jenkins, 2009). The studies reviewed in the next section describe the nature of feedback and its impact on students learning and motivation.

Higgins, Heartley, & Skelton (2002) conducted a study to explore the use of written feedback as a tool to assess students' understandings during an interaction between teacher and students in a community of practice and how students react to teacher's written comments on written assignments. The purpose of this study was to document whether, with potential barriers and confusing language of assessment feedback, students would disregard the use of feedback. Aspects of the constructivist theory of learning were used to encourage deep learning in students. Data were collected from 19 higher education students from two different institutions using a semi-structured interview and a Likert scale questionnaire. No information was given on how these instruments were validated. Results indicated that students prefer feedback that provides information on strengths and weaknesses and a means for improvement, as well as focusing on achieving grades alongside intrinsic motivation. Higgins et al. (2002) identified structural barriers to feedback involving quality and quantity and suggested that future studies needed to address these issues.

Jenkins (2009), in a similar study, created a multifaceted formative assessment to provide students with feedback and gives them the opportunity to act on the directives before re-submitting the revised version. In this action research, she collected data on written feedback through soliciting and validating students' ideas from questions posted in an open comment section of the e-learning platform. Analysis of students' perception of the assessment process, indicated that students agreed that assessment feedback helped them with learning. Jenkins (2009) suggested that formative assessment practices can be encouraged if used as a replacement for unnecessary summative assessment, so long as teachers do not see it as an additional burden. This, she asserts, will improve the opportunity of feedback fundamental for more in-depth learning experiences.

Kang et al. (2014) conducted a mixed-method study to examine ways in which teachers provided students with feedback as written scaffolds in assessment tasks and how it impacted students' construction of written explanation. The rationale for using the instructional scaffold was that formative assessment as a scaffold moves learning forward within the zone of proximal development. The purpose of this study was to see how and why particular forms of scaffolds embedded in assessment tasks guide students' construction of written evidence-based explanations. Data were collected from 76 assessment tasks designed by 33 first-year teachers. Five different types of scaffolding techniques were identified from these tasks: 1- contextualized phenomena, 2- rubrics, 3- checklists, 4- Sentence frames/starters, and 5- drawing and writing (Kang et al., 2014). Analysis of students' written responses for evidence-based explanations relative to science formative interactions indicated that effective scaffolding created an opportunity for students to demonstrate scientific understanding. Also, the quality and combination of scaffolds used were more effective than the number.

The articles reviewed in this section discussed the need to provide appropriate feedback to students and scaffold them to close the gap between the current and expected level of understanding, structural barriers that hindered these processes, and features involved in designing and giving support. Higgins et al. (2002) and Jenkins (2009) both used written feedback to understand how students interact with different components of formative assessment in a community of practice. They concluded that multiple criteria of judging performance exist; feedback should clarify the aim and objective of formative assessment to reveal a range of conceptions captured in classroom conversations. They suggested that teachers' reflective and directive written feedback could satisfy students' need for both gaining higher conceptual understanding and higher achievement. Jenkins (2009) and Kang et al. (2014) studies demonstrated the need for explicit feedback that provides contextualized scaffolds to help students construct evidence-rich explanations. These studies add to the quality of formative assessment. However, the process of providing feedback and scaffolding needed restructuring to be enticing to both teacher and student. That is, teachers need the ability to provide reflective and corrective feedback that will scaffold and give students the confidence to complete assessment tasks.

Formative Assessment Support and Change in Teachers Practices

Teacher change is possible, but slow. Situational factors such as the nature of the reform message, support (resources), guidance, collaboration, belief in teaching and learning, and timing influence teachers' transformation (Borko et a., 2000). Teachers made decisions on how to process the reform message depending on whether the message communicated initiated stress in their conceptual understanding (Ebert & Crippen, 2010). They also made sense of reforms collectively on shared understandings of its message as they interpreted and created their responses to students' ideas (Furtak, 2014). Evidence supports that teachers want to change their assessment practices to reflect those advocated by the reform message so long as it reproduces features of teaching and learning mandated for K-12 classrooms (Borko et al., 2000). However, teachers are reluctant to disrupt the routine established in their classrooms because embracing the practice of formative assessment demand reconstructing the teaching practices that they have worked so hard to build and see it as successful (Lee & Wiliam, 2005).

Sato, Chung, and Darling-Hammond (2008) explored how the National Board Certification process as a professional development learning opportunity can improve mathematics and science teachers' everyday classroom assessment practices. Two groups of self-selected teachers were recruited and data from numerous sources collected from three experimental or National Board Certification candidates and three regular teachers. A 5-point indicator scale rubric on each of the six assessment measures was used for data analysis. The findings showed an improvement in formative assessment practices for National Board candidates better than for non-National Board teachers. The result is evident for "their conceptions of assessment as shifting from a focus on grading to the use of assessment for formative purposes" (Sato et al., 2008, p. 23). However, the authors indicated a noticeable improvement only to a small sample, although increased in professional development for non-National Board teachers lead to improved formative assessment practices. They suggested repeating this study for future research.

Lee and Wiliam (2005) conducted a case study to describe the process of teacher change and the development of formative assessment practices that foster this change. Twenty-four teachers participated in a support program to develop expertise in assessment for learning. The teachers practiced with their students before implementation because Lee and Wiliam (2005) believed the practice of formative assessment demand that the teachers change their old ways of doing things with the students. Initial data on teachers' views and beliefs were collected through interviews and observations on teachers' implementation of formative assessment. Two teachers were studied closely to obtain a detailed account of changes in teachers' practices. Analysis of this data indicated that Lee and Wiliam (2005) identified six common factors that could be attributed to the significant changes in teachers' practices. (1) Credible evidence that motivated teachers to change their practices. (2) Having practical ideas to implement in the classrooms immediately. (3) Continuous support from the researchers and professional learning community. (4) Interventions to provide opportunities for reflection on immediate actions and further perspectives and insights. (5) Enough time to support teachers' slow pace of change (6) Flexibility to use as many strategies presented to develop their formative assessment practice (p. 13). Lee and Wiliam's (2005) suggestion for the future was '*what works approach,'* a model robust in professional development learning, but also flexible enough to consider teachers' differences, capable of generating greatest effect across all teachers.

Phelan, Choi, Vendlinski, Baker, and Herman (2011) conducted a similar randomized study to see whether using a formative assessment strategy would increase students’ performance on the assessment of big math ideas relative to the performance of a comparative group. All students took a pretest and a transfer outcome test. The treatment students received regular checkpoint understanding as formative assessments while the treatment students were provided with intervention to enhance understanding of big ideas plus additional resources. The result indicated that treatment students with higher pretest results had higher transfer outcome result or benefited more from intervention. Treatment students outperformed control students on the test. However, the intervention had more impact on most difficult mathematic concepts. Phelan et al. (2011) suggested future studies to explore students’ growth trajectory and the variability of test used.

Another professional development intervention study conducted by Randel, Apthorp, Beasley, Clark, and Wang (2016), was to estimate the impact of classroom assessment on students' mathematics achievement and students' involvement, and on teacher's knowledge of classroom assessment and assessment practice. The sample comprised of 64 teachers, 32 fourth grade control group, plus 32 fifth grade experimental group. Randel et al. (2016) hypothesized that students will benefit from more explicit learning goals, better assessment, and feedback. Also, improved teacher assessment knowledge and skills would positively impact students' achievement in mathematics. Data were collected from both cohorts using the same Colorado state 3rd grade test for pretest, but cohort 1 implemented 5th grade test for posttest while cohort 2 implemented 4th grade test version. The instruments were survey of teachers' background characteristics, Classroom Assessment of Student Learning (CASL) implementation log, and students' achievement log. The results indicated that use of CASL with fidelity may result in positive impact for test group than control group teachers, as well as for controlled outcomes like teacher knowledge and involvement of students in assessment. However, little impact was evident on outcomes beyond teacher's control like performance on state test. Moderate degree of implementation suggested that it is easy to improve teacher assessment knowledge and students' involvement than to practice or provide descriptive feedback to students.

Meusen-Beekman, Binke, and Boshuizen (2016) used a formative assessment (peer and self-assessments) intervention on six grade students to explore its impact on students' self-regulation, motivation and self-efficacy and whether there is a difference between these different forms of assessments. They assigned a total of 695 students into peer-assessment, self-assessment, and control group for a 27 weeks' classroom intervention. Data sources included self-assessment questionnaires and interviews. All students took a pre- and posttest assessment.

Multilevel analysis indicated that the used of self and peer-assessment were effective in developing self-regulation and motivation in treatment students. No significant difference observed between peer and self-assessment interventions, and students' self-efficacy was not affected by the intervention. Students indicated that their attitudes towards formative assessment became more positive, and they felt more confident giving and receiving feedback. According to Meusen-Beekman et al. (2016), the result should be received with caution as it arose from self-reported data. However, they suggested that the study is important to help primary school students adjust and use learning strategies to motivate and improve achievement. Future study is encouraged to determine whether planning or providing feedback was a better intervention.

A study to examine the effects of changes in twenty-two fourth-grade mathematics teachers' formative assessment practices, after professional development reviled a similar pattern on students' achievement (Andersson & Palm, 2017a) to Meusen-Beekman et al. (2016) study. Data were collected from two randomly selected groups of teachers' students' pre- and posttest scores. The result from a one-way between Analysis of Covariance (ANCOVA) indicated that students in the classrooms of teachers who participated in the professional development outperformed their counterparts in the control classrooms in a posttest score. Anderson and Palm asserted that this study provided empirical evidence for the impact of formative assessment on students' achievement and that professional development on teachers' implementation of classroom assessment significantly impacted student learning. They suggested that future professional development should have provision for continuous collaboration for participant teachers, provide ways to implement a formative assessment to diverse students of mixed abilities and behavior, and for an extended duration.

Andersson and Palm (2017b), in a, follow up study, explored the characteristics of changes in classroom practice from a combination of formative assessment strategies and the link between the characteristic of these changes and learning opportunities for students. Data were collected from twenty-two fourth-grade teachers through interviews and observation of classroom practices. Analysis of this data revealed the complexity of formative assessment practices. It showed changes ranging from enhancing existing strategy focused on “big idea” with new activities, to completely changing old practices with new ones. Andersson and Pam (2017b) suggested that teachers would have to make a drastic change in their practices to implement an effective formative assessment.

Gearhart et al. (2006) conducted a study on teachers' evolving expertise in interpreting students to work with the help of portfolios and to discover needed resources and teachers' challenges using weak assessment tasks and criteria. The goal was to document the changing relationships between teachers' practices and their purposes. The assumption was that teachers develop expertise with the interpretation of student work and associated assessment concepts with the repeated alignment of old and new understandings and practices. Data were collected from three experienced teachers using teachers' interviews and portfolios. Analysis of all three cases indicated that teachers slowly embraced improving their interpretation of student work through integrating new assessment concepts. Gearhart et al. (2006) suggested a need for both qualities embedded formative assessment resources and the development of teacher assessment expertise. As well as future research based on teachers' assessment system and on grading, informal assessment, and designs for unit assessments.

Successful education reform is predicated not only on teachers’ understanding, participation, and support but also in their views. Yan and Cheng (2015) explored the relationship between primary school teachers' attitudes, intentions, and practices regarding formative assessment. Their study is framed under the Theory of Planned Behavior. Survey data were collected from 10 teachers and analyzed using Rasch scale followed by path analysis. The outcome was that the teacher's intention to practice formative assessment was influenced by 1) instructional attitude, 2) subjective norms, and 3) self-efficacy. The implication for this study was that teachers need a positive instructional attitude and high self-efficacy for teachers' intention towards formative assessment. To change assessment culture in the classroom, teachers needed to change their conceptions of assessment and embrace the intention to change.

Studies in this section provided insight into factors that influenced teachers' intention to practice or change their practices towards formative assessment practices and the impact on students' achievement. The interventions highlighted the complexity in teachers' formative assessment practices. Some teachers exhibited a slow change from assigning grads to a formative purpose (Lee & Wiliam, 2005; Sato et al., 2008). Some enhanced their practices by integrating reform strategies (Andersson & Palm, 2017b; Gearhart et al., 2006), and some changed their instructional attitudes (Yan & Cheng, 2015). The interventions also improved students' achievements (Andersson & Palm, 2017a; Lee & Wiliam, 2005; Phlelan et al., 2011; Randel et al., 2016; Sato et al., 2008). In some cases, treatment students with higher pretest scores improved more than those with low pretest scores (Phlelan et al., 2011) and in others, higher pretest scores improved students' attitudes towards formative assessment (Meusen-Beekman et al., 2016). These studies suggested future research with flexibility in implementation, provision for continuous collaboration, planning and providing feedback, accommodation for diverse students and teachers, and improved design for assessment.

Formative Assessment and Teacher Knowledge

Formative assessment is grounded in the concept of Modern Validity Theory (which assumes that the validity of the test is second to the inference or interpretation is drawn from it and its use), on the quality of assessment measures, and on the quality of assessment process (Herman, Osmundson, & Silver, 2010). The quality of the formative assessment process specifies clear learning goals, iterative process of eliciting, interpretation, and use of evidence. These elements serve as inference for the next step in teaching and learning to reduce the gap about the goal (Herman et al., 2010). Herman & Choi (2008) suggested that the quality of interpretation or accuracy of the decision-making process is essential. Thus, the validity of formative assessment rests on the teacher's appropriate interpretation and use of results. Similarly, teacher's pedagogic content knowledge is closely related to teacher's formative assessment practice and practicing formative assessment builds teacher's pedagogic content knowledge (Furtak, 2012). The studies in this section explored efforts to address teachers' limited knowledge and difficulty implementing formative assessment (Buck, Trauth, & Kaftan, 2010).

Herman et al. (2010) examined science teachers' measures of formative assessment practice using data from implementation, and the effects of adding curriculum embedded measures to a hands-on science program for upper elementary school students. Data on teachers' assessment practices were collected on 39 teachers using a Teacher-Content-Pedagogical-Knowledge instrument, content survey, and teacher self-report. Data from observations and interviews were collected only from a small sample. Results between these constructs showed: (a) no relationship between teachers' self-report of their content-pedagogical knowledge and direct demonstrations of such knowledge; (b) no relationship between content knowledge and ability to analyze and suggest next step for instruction based on students' responses; (c) no relationship between establishing goals and analyzing students' work toward those goals (Herman et al., 2010). However, teachers who reported establishing and communicating their learning goals, also reported coordinating their assessments with those goals. The teachers who reported aligning their goals and assessment also reported that they analyzed student and group work and used a variety of strategies to assess student understanding. Herman et al. (2010) suggested developing valid measures of assessment practice for accurate research findings and more training for teachers, especially preservice teachers.

Buck et al. (2010) conducted an action research to explore the reconceptualization efforts in preparing preservice teachers, to guide the inquiry process with formative assessment and to use the understanding to improve teacher preparation program. They employed an implicit and explicit method of content delivery and compared which one improve preservice teachers' knowledge of formative assessment. Their ongoing and iterative data collection and analysis process were based on four analysis criteria. (1) understands the purpose of formative assessment, (2) relates formative assessment to students' conceptual development, (3) links formative assessment outcomes to instructional planning, and (4) demonstrates an understanding of relational processes inherent to formative assessment. Findings indicated that preservice teachers could successfully embed formative assessment into their lessons. However, more than half still demonstrated a naïve understanding in two areas; relating formative assessment to students' conceptual development and demonstrating an understanding of relational processes inherent to formative assessment. Buck et al. (2010) noted that explicit and conceptualization approaches to formative assessment in the method course influenced preservice teachers' construction of a deeper understanding of formative assessment than the implicit approach. In addition, preservice teachers were unable to transfer their working conceptions of formative assessment to other appropriate pedagogical strategies, had limited understanding of the collaborative nature of formative assessment, and the role of students in formative assessment (Buck et al., 2010).

Herman, Osmundson, Ayala, Schneider, and Timms (2006) conducted a study to describe the quality of teacher assessment practices and its impact on students' learning and to formulate implications for professional development and for future research. They used multiple measures of teachers' knowledge, instructional practices, and student learning to collect data. They used a quantitative and a qualitative method of data analysis. The overall quality of formative practices of the eight teachers observed was judged to be at the beginning stages of effective formative practices. Results of the Hierarchical Linear Model HLM analysis indicated that ways in which teachers use quality assessment tools to support and promote student learning and achievement were insignificant. It should be noted that Herman et al. (2006) acknowledged the limitation of the HLM due to the small sample size. The implication of this study was the need for teachers to improve the quality and quantity of their assessment practices, allocate time to design and teach new curriculum and assessment, and provide timely and scaffold feedback.

Herman and Choi (2008) conducted a similar study to explore the accuracy of teachers' judgment of student learning and its relationship to students' performance. Data were collected from thirteen teachers using a pre-test and a post-test, in addition to the measure of students' conceptual understanding of buoyancy and students' developmental responses to reflective lessons. Herman and Choi's (2008) hypothesized that "teachers' accuracy in judging student performance is positively associated with subsequent student learning" (p. 15). However, they acknowledged that the measures for rating teachers' interpretation were imperfect because the teachers were using gross estimates versus the centralized scores they were using. Their findings indicated a positive impact of formative assessment and that teacher's accurate judgment on student learning levels parallel to their ability to improve student learning. It also revealed challenges in assuring accuracy in teachers' interpretation, a necessary precursor to providing useful feedback, and maximizing advancement in instruction and learning (Herman & Choi, 2008). Suggestion for future studies was towards more focus on assessment accuracy and to foster the conditions in place and characteristics of good practice.

Heritage, Kim, Vendlinskin, and Herman (2008) conducted a generalizability study to know whether adapting instruction to meet students’ needs is a link to a teacher’s competence. The purpose was to determine the component of variability in 118 six-grade mathematics teachers’ knowledge that is most likely responsible for total scores on the teacher knowledge scale, and if this finding can be applicable to teachers in general. The teachers completed a series of assessment tasks rated by experts on a 4-point scale rubric to assess teachers’ pedagogical knowledge of mathematics concepts. The results indicated that teachers’ scores were not generalizable with regards to identifying key principles, evaluating student understanding, and formulating the next step in instruction. They asserted that teachers were better at making inferences about student’s level of understanding from assessment information than deciding the next step for instruction. Heritage et al. (2008) suggested that future studies should explore teachers’ ability to formulate the next step and how to adapt their instruction.

In exploring science teachers’ PCK for inquiry and formative assessment practices, Buck et al. (2010), Herman and Choie (2008), and Herman et al. (2006, 2010), suggested that most teachers had difficulty interpreting students’ responses and did not use assessment results to adjust instruction. However, it was noted that teacher accuracy of interpretation of student work reflected achievement gains. Buck et al. (2010) determined that explicitly teaching formative assessment improved preservice teachers' knowledge, but these teachers had difficulty interpreting students' ideas and using the results. An interesting observation is that there is no relationship between teacher knowledge of formative assessment and their ability to practice (Herman, 2010). Even where there is a relationship, the teacher had difficulty transferring the knowledge to other domains or to adjust instruction (Buck, Trauth-Nare, & Kaftan, 2010). Furthermore, such successes could not be generalized to other teachers (Heritage et al., 2008). A consensus among these studies was the need for more training in quality assessment- accurate interpretation and, most of all, formulating the next step of instruction (Herman & Choie, 2008; Herman et al., 2006).

Sequencing Learning for Formative Assessment Practice

The learning sequence is a content-specific practice that can help to understand students' ideas of a domain as well as a representation of how students' ideas develop towards "more sophisticated thinking" (Furtak & Heredia, 2014, p. 4). It is built on the logical development of scientific concepts. Its recognition as a promising tool to foster teachers' formative assessment practices is increasing because it can assist the teacher in identifying and making inferences of evidence gathered from students thinking (Furtak, 2012). The teacher thus can use a learning sequence to understand student learning trajectories, predict areas of weaknesses, and plan a formative assessment to address the different concepts (Furtak & Heredia, 2014). This is possible given that learning sequence can represent multiple trajectories of learning (NRC, 2007).

Furtak (2012) reported a study on six high school Biology teachers using a learning sequence for natural selection to enhance their practices of formative assessment. Data was collected via interviews and videotapes of teacher classroom assessment conversations. An analytical coding of videotapes for ideas students share during discussions and inferences teachers make about students' ideas were developed, as well as descriptive coding for how teachers use learning sequences to inform interactions and make inferences about students thinking. The results indicated that learning sequences helped teachers create curricular units structured on the logical development of scientific concepts and helped them identify students' misconceptions but did not help them act on students' ideas and adjust instruction. This study, according to Furtak (2012), shows the complexity of students' ideas shared during the formative assessment and how students' ideas may be distributed across the learning sequence. She suggested that the process of sequencing lessons needed additional support to help teachers adapt their instruction.

Furtak and Heredia (2014) conducted a multiple case study to explore how learning sequence acts as a boundary object to coordinate the work of two communities of biology teachers in making instructional plans, developing formative assessment, and interpreting student ideas. Analysis of professional development session videotapes and teacher interviews indicated that "learning sequence took on different meaning through its use at each of the different schools and served the purposes of planning instruction, developing formative assessments, and interpreting student ideas in different ways" (Furtak & Heredia, 2014, p. 32). The ability of the learning sequence to support teaching and learning depends on the context used by teachers- as a tool or process. They concluded that soliciting teachers' input in co-constructing learning sequences can encourage teacher buy-in and help them recognize insight of theirs and students' ideas. The importance of the result to NGSS, according to Furtak and Heredia (2014), is that when teachers participate in creating learning sequences, they can easily adapt it to their own context and standard. Further studies are needed to understand how the representation of student ideas in a learning sequence might help guide teachers' instructional practices.

In a similar focus on learning sequence, Furtak et al. (2016) conducted a three-year longitudinal study to explore formative assessment abilities of purposefully selected nine tenth-grade Biology teachers in relation to their students' learning. They facilitated a Formative Assessment Design Cycle (FADC) meetings where teachers sequenced the curriculum and identified areas that needed further instruction, constructed formative assessment tasks, practiced using the activities with each other and reflected on, and enacted the tasks in their classrooms. Data collection occurred through the interpretation of students' ideas in relation to learning sequence, a product of formative task, a videotape of enactment, and pre-and posttest of students' achievement. Furtak et al. (2016) analyzed the data using HLM and ANCOVA, respectively, because of the nested and descriptive nature of the data. The results indicated that task design from base year to year three was significant; asking questions to elicit students' ideas was significant but it did not have positive impact on students' achievement; Interpretation of students' ideas was statistically significant; the quality of teachers' responses to students' ideas with respect to learning sequence was significant. They proposed that teachers need to provide students with eliciting questions and feedback to alternate between the dialogic and authoritative function of discourse. They inferred that learning sequence might have influenced ways teachers interpreted students' ideas. In addition, a scaffold of learning sequence in each domain could be a scaffold of the teacher's interpretation of student's ideas. Furtak et al. (2016) suggestion for future research is to repeat this study using a two-group experimental design and to investigate the relationship between teacher assessment task and teacher formative assessment practice.

Three-Dimensional Teaching and Formative Assessment

The Framework of K-12 Science Education (NRC, 2012) and the Next Generation Science Standard (NGSS Lead State, 2013) reform encourages a shift in focus from assessing science ideas to assessing how students figure out phenomena and construct solutions to problems. This reform way of learning requires students to integrate the three dimensions that define science literacy (Reiser et al., 2017). All students can actively engage in using and applying knowledge of the discipline to promote deep learning. There is overabundance of research on two-dimensional teaching and learning (Disciplinary Core Ides (DCI) and Science and Engineering Practices (SEP)), but the existing literature on three-dimensional teaching and learning (DCI, SEP, and Crosscutting Concepts (CCC)) is theoretical with no clear application to classroom practice or research (Fick, 2017; Fick & Songer, 2017; Harris et al., 2015). Researchers suggest that CCC can serve as a lens to analyze phenomena, as a bridge to connect concepts across disciplines, as a tool to understand content in different ways depending on the purpose and nature of inquiry, and as a rule to guide the use of DCI (Rivet, Weiser, Lyu, Li, & Rojas-Perilla, 2016). Task created as formative assessment provides evidence of integrated science knowledge. According to NRC (2014), tasks should compose of multiple components to reflect the interconnectedness of the practices, reflect progressive learning by soliciting information along a continuum of results in each grade and be interpretive to evaluate a range of student responses and to guide instructional next step for teachers. This section described studies at the forefront of the 3D integration movement in an attempted to shed light on this new way of teaching and learning science. The articles focused on (a) student-centered three-dimensional teaching practices; (b) design approaches to support teachers in three-dimensional learning; (c) integrated assessment items for assessing students' three-dimensional learning.

Student-centered three-dimensional teaching practices

Fick (2017) piloted an experimental study design to test the impact of integrating CCC into an instructional unit. The goal of the study was to assist the student in learning to use the

CCC framework to clarify misunderstanding, ask questions of new phenomena, and make the connection of science ideas across context. Fick (2017) employed a simplifying scaffold for students as a CCC framework to support the examination of phenomena. The author collaborated with one teacher to develop and revise the curriculum, and the teacher enacted them. Data were collected from video and audio recordings and from pre-and post-test conceptual models. The analysis was done using descriptive codes of the dimensions and rubrics. The author reported that the teacher used students' conceptual models and classroom dialogue to demonstrate how CCC provided students with the opportunity to learn. The teacher used CCC to frame classroom activities to frame discussions that developed a student's understanding of CCC as a lens to examine phenomena and serve as a component of student's conceptual model that highlights their understanding. The implication for their study illuminated the role of the learning environment in supporting student's three-dimensional learning. Further research is needed to see whether students could use their understanding of CCC in one context and apply it to another or make connections between ideas. Fick suggested that teachers can use the Framework as a tool/lens before using it as a bridge to support the student to see the purpose of the Framework. Another suggestion was for teachers to have students apply the Framework across science topics to deepen their understanding of a new concept. Fick made a bold claim that the study has the potential to support students and teachers to engage in 3D science learning.

Lauren, Lutz, Wallon, and Hug (2016) conducted a descriptive pilot study to examine how a collaborative board game about honeybees that simulate worker bees within a colony, could be used to integrate the three dimensions of science education. Disciplinary Core Ideas (DCI) was represented by core ideas of "social interaction," "group behaviors," and "variation of traits." Crosscutting Concepts (CCC) were represented by the concepts of "cause and effect" and "system and system models" (the hive). Science and Engineering Practices (SEP) was represented by limitations and approximations of the model, considering scientific evidence, generating data, making predictions, and evaluating the model. The game provided a means for students to incorporate scientific evidence on how genetic and environmental factors influence variation of traits and social behavior and communicated understanding and strategies. Students also evaluated the game as a model in an authentic classroom setting. Lauren et al. (2016) suggested that games and simulations if accompanied by collaboration, can be a promising way to engage students in 3D learning. They acknowledged that most teachers struggled alone on how to align their lesson to NGSS, so teachers and educators need resources that are engaging, and student driven. Given that the game incorporated many aspects of scientific practices, Lauren et al. (2016) suggested that teachers might consider providing a rubric to guide the discussion and evaluation of the game, as well as an extension activity to challenge students thinking in the future iterations of the game.

Jasti, Lauren, Wallon, and Hug (2016) conducted a similar board game that teaches students about the biomagnification of toxicants across trophic levels while engaging in 3D learning. The authors recruited five teachers who attended a summer professional institute to learn the skills for 3D learning. The teachers enacted the Bio Bay activity in their classrooms, and the primary investigator collected data through observation and two in-depth individual interviews. The authors indicated that teachers did not all use the Bio game the same way but used it mainly to make connections to concepts in the DCI and to practice modeling and data collection. Teachers also indicated that the Bio Bay activity provided students the opportunity to engage with a real-life phenomenon as a meaningful 3D learning experience. In addition, students had the opportunity to generate, analyze and interpret data (visually using graphs), make the connection to the real world, and evaluate the game and model (Jasti et al., 2016).

Harris et al. (2015) conducted a study on middle school curriculum materials called Project-Based Inquiry Science. The material was designed to engage students in science practices of constructing explanations and developing and using models to demonstrate their understanding of disciplinary core ideas in Earth and Physical science. A randomized control trial was conducted with two groups of six grade students in science classrooms across 42 schools in a district. Two teachers per school were recruited, some schools implement the project-based science curriculum and others the district adopted curriculum. Both groups received training on the new Framework and enacted their respective curriculum. Harris et al. (2015) wanted to know (1) the extent to which the project-based curriculum can be implemented with fidelity; (2) the extent to which fidelity of implementation is related to district support and prior years' assessment levels; (3) the impact of project-based curriculum on student science learning; (4) how this curriculum impact students of different backgrounds. The authors developed multi-component assessment tasks to assess student understanding of disciplinary core ideas. They reported that students in the treatment group outperformed their counterparts in measures aligned to core science ideas and science practices. They also found that classrooms with low achieving students benefited more from the project-based curriculum, and fidelity of implementation was not related to the concentration of low achievers. They asserted that curriculum materials that incorporated science practices along with disciplinary core ideas, can foster students' three-dimensional learning. They suggested that providing teachers with reform-based curricular material, district involvement, and support for its implementation is necessary for realizing the Framework's vision and key principles in a district. Limitation of the study concerns random selection, which provided a special condition that is difficult to replicate, the treatment teachers received a higher support than the control teachers and more than in their everyday instruction. Also, it was difficult to rate the fidelity of implementation using implementation logs, and the project-based curriculum was not truly aligned to the Framework's vision. Harris et al. (2015) claimed that the study offered early effort to measure student's integrated science proficiency, provided the basis for the need for the new curriculum plus new assessment aligned with the Framework and as found in the NGSS. The study also provided evidence for the role of assessment in promoting equity.

Design approaches to support teachers in three-dimensional learning

Reiser et al. (2017) conducted a study based on a two-part program for scaling 3D science professional development where teacher leaders develop expertise in 3D learning. The leaders, in turn, facilitated the study groups of teachers in 3D science activities, analyzed student learning, and investigated classroom interactions. The authors recruited and trained 24 teacher leaders as experts in 3D learning to facilitate 420 teachers divided into 22 study groups. The purpose of the study was to answer three questions. How does professional development focus on classroom practice help teachers improve proficiency with 3D science? How do professional development focus on classroom practice influence teachers' confidence and beliefs about learning and teaching consistent with 3D learning? How does professional development focus on classroom practice help teachers develop PCK needed to support 3D learning? (Reiser et al., 2017, p. 285). Data were collected through a pre- and post-professional development survey and analyzed. Reiser et al. (2017) reported that teachers became more proficient in using disciplinary ideas from domain to explain phenomena; Teacher's confidence and feelings of readiness to take on challenges of the new reform improved throughout the professional development; Overall, the professional development improved sophistication of teacher reasoning about pedagogical scenarios involving practices; Teachers showed better understanding and facility generating situations in which models were developed as generative tools for students to construct, argue for, evaluate, and revise explanations (p. 294). However, Reiser et al. (2017) cautioned that their findings were suggestive because the feeling of having confidence and being prepared does not mean the teacher is capable. Also, teacher attitude can influence future participation. The limitation of their study is the difficulty of exploring whether and how an increase in teacher's expertise leads to changes in classroom interaction and subsequent student learning. They suggested that further research is needed to examine study group interaction for learning that is most lucrative in helping teachers grapple with complex questions of practice. They also recommended studies that consider the facilitator's strategy that is effective in leading study groups and how to support these strategies (Reiser et al., 2017).

Another study looked at preservice teachers' interaction with 3D learning in preservice teachers. Richmond, Parker, and Kaldaras (2016) examined explanations constructed by teacher candidates as scientific practice for supporting student's 3D learning. The authors recruited a stratified sample of nine teacher candidates (three Chemistry & six Biology majors) in their final two years of the preparatory program. Teacher candidates were taught in their method course to organize explanations into what/how/why framework. Data were collected from warmups and lesson plans and analyzed using a modified constant comparative and an inductive multi-stage approach. The authors examined teacher candidate's ability to organize explanations in this framework according to what happens to a case (data or observation), how things happen (patterns or laws), and why or casual explanation of model or pattern. They found that teacher candidate's ability to articulate complete and accurate casual explanation for phenomena exist along a continuum. A teacher candidate with an explanation at the upper end on the continuum could provide explanations without support from the standard, while a candidate with an explanation at the lower continuum struggled even with support from the standard. Richmond et al. (2016) documented that teacher candidate's ability to construct complete and accurate explanation is not related to course performance or major; the teacher candidate must be able to provide accurate account of a phenomenon to be able to develop casual explanation; candidates who came in with explanatory skills continued to display them, and those without struggled; candidates who struggled to provide causal explanation of a phenomenon lacked deep understanding of CCC. The authors suggested that teacher candidates need specific and ongoing support to help them structure scientific explanations around CCC, essential for 3D science teaching and learning. Their implication is that the language of the framework could be used to characterize the type of scientific explanation provided by teacher candidates and inform the work of future teachers. It could also provide stakeholders with a common language for discussing scientific explanations for 3D learning (Richmond et al., 2016). They proposed that future studies should provide an explicit scaffold for developing skills to recognize key principles driving processes/systems to help teacher candidate frame explanations for phenomena. The limitation of their study was the small sample size that hinders the ability to make a correlation between the degree of explanation and major and use of a variety of topics.

Integrated assessment items for assessing students' three-dimensional learning

Fick and Songer's (2017) study described the nature of alternative integrated science knowledge demonstrated by students in response to an integrated science assessment task. The purpose of the study was to answer one question, "what alternate integrated science knowledge do eighth-grade students demonstrate in response to integrated assessment items?" (Fick &

Songer, p. 140). Data were collected from six eighth grade charter middle school students using a continuous screenshot activity on written and audio responses of 19 assessment questions. The study focused on characterizing the type of knowledge student demonstrates pre-instruction relative to assessment prompts. The assessment task gave students the opportunity to analyze data, find patterns, and make predictions about the cause and effect of climate change. The results showed examples of assessment items and students' responses that represent the students' progress on integrated science knowledge. The task revealed many levels of alternative integrated science knowledge held by students. Integrated assessment can provide insight on students' struggles coordinating science content and practices for three-dimensional learning and display a continuum to compare student's alternate science knowledge. Fick and Songer claimed that their work could be used for developing teaching strategies to support teachers' development of integrated science knowledge; Students integrated responses can reveal the challenges they face if their explanations are incomplete or inaccurate. The authors proposed that there is a need for research that illuminate's students' misconceptions of integrated science tasks and that highlights intermediate and advanced forms of integrated science knowledge.

Studies that explored integrated science learning provide a new perspective on how science is done in classrooms to reflect the K-12 Framework's vision. Those that focus on curriculum materials reported that it had an impact on classroom practice and on student learning (Fick, 2016; Harris et al., 2015; Jasti et al., 2016; Lauren et al., 2016). Creating instruction material like the board games and simulations with real-life phenomena that integrate the three-dimensions with a collaborative component, are promising ways to engage students in 3D learning (Jasti et al., 2016; Lauren et al., 2016). Three-dimensional oriented curriculum materials benefit low achievers more, and 3D teaching can level the playing field for all students (Harris et al., 2015; Lauren et al., 2016). These studies also reveal that just providing teachers with materials on how to carry out 3D instruction is not enough; it should be accompanied by resources and support from the district and administrators. In addition, a supportive context can alleviate accountability pressure and foster buy-in (Harris et al., 2015). Studies that provided professional development to support teacher 3D learning indicated that it provided a practical approach and improved teacher's reasoning about the phenomenon and developing models (Reiser et al., 2017). The training helped teachers structure scientific explanations around CCC and improved the ability to articulate complete and accurate explanations (Richmond et al., 2016). Professional development enhances teachers' understanding of and tendency to enact 3D learning and scaffolds teachers to develop skills to model constructing scientific explanations to students. However, the teacher cannot become proficient in 3D learning without practice (Reiser et al., 2017; Richmond et al., 2016). From these studies, we see that assessment tasks for integrated science practices must be multicomponent to assess the different elements of each practice and measure student's integrated science proficiency (Ficks & Songer, 2017; Harris et al., 2015). The what, why, and how Framework can serve as an analytical and descriptive tool to represent explanations of teacher's 3D learning. The CCC framework can be used as a common language to characterize types of student explanations using multicomponent (Richmond et al., 2017) assessment tasks. Fick and Songer (2017), Reiser et al. (2017), and Richmond et al. (2016) all demonstrated how integrating practices can move student learning along a continuum from an emerging to a distinguish learner. These studies additionally highlighted the role of the learning environment in supporting students' 3D learning.

Significance Research Findings

The findings from these articles were grouped into three categories, teachers' role and thinking with practice of formative assessment, students' role and achievement, and the status of three-dimensional teaching and learning.

Teacher role and thinking with formative assessment

Teachers decide on how to process the reform message or intervention and adopt or reject it depending on the message communicated, whether it initiates stress in their conceptual understanding or not. A teacher who implicates self in the reform message or intervention processes, and continues to tinker with information, realizes the conceptual change in their classroom practice. Meanwhile, a teacher who engages in "benign positive appraisal… I already know this, and it is not for me" (Ebert & Crippen, 2010, p. 376), rejects the reform message. Results from this review indicated that in the process of practicing FA, teachers changed their role from a closed authoritarian to an open dialogic behavior and developed a deeper conceptual understanding of FA (Buck & Trauth-Nare, 2009; Buck et al., 2010; Furtak & Heredia, 2014). Others changed their thinking of FA from evaluating to diagnosing students' understanding of the concept (Offerdahl & Tomanet, 2011; Sato et al., 2008). Teacher's praxis changed with the used of various interventions such as with the use of rubrics, notebooks, personal response systems, lesson sequences, and portfolio (Anderson et al., 2007; Aschbacher & Alonzo, 2006; Feldman & Capobianco, 2008; Furtak & Heredia, 2014; Gearhart et al., 2006). The change was also documented with professional development, which, according to evidence, must provide enough continuous support, be reflective, have practical ideas and credible evidence, and be flexible (Morrissette, 2011; Sato et al., 2008; Wiliam et al., 2004). However, these changes in teachers' practices were reported to be slow and brief and differed from one teacher to another. Their experiences might range from no change to a complete shift from novice to expert (Feldman & Capobianco, 2008). Evidence supports teachers' willingness to change their practices to reflect those advocated by the reform message or intervention so long as it reproduces features of teaching and learning mandated for K-12 classrooms (Borko et al., 2000). However, teachers are reluctant to disrupt the routine established in their classrooms since embracing a new practice demands reconstructing the teaching practices that they have worked so hard to build and perceived as successful (Lee & Wiliam, 2005). Another factor that might have influenced teachers' thinking, role and change of FA practices is limited knowledge and skills of FA practice (Aschbacher & Alonzo, 2006; Herman et al., 2006; Herman et al., 2010; Hickey et al., 2012).

Student role in FA and student achievement

Students play an important role in the success of formative assessment. Social interaction mandates that both teacher and student share responsibility, the teacher designs, implements the task, and guides students by recognizing students' knowledge and experiences while students take active role in the learning within the environment (Windschitl, 1999). When teachers provide students with descriptive feedback and allot them time and opportunity to revise their work, it encourages students to demonstrate their improved abilities, especially if the comment comes before grades (Higgins et al., 2002; Jenkins, 2010; Kang et al., 2014). Formative assessment feedback that motivates students to do better also improves their conceptual understanding and achievement on an exam (Hickey et al., 2012; Wiliam et al., 2004). The role and nature of group discourse couple with teacher intervention affect the quality of students' performance (Anderson et al., 2007). Wiliam et al. (2004) claimed a positive impact on students' achievement with formative assessment; however, no data was provided to corroborate their findings, and their data comparison method were flawed. Conversely, Yin et al. (2008) reported no change in students' achievement, conceptual change, and motivation. Students' achievements were also reported in cases where teachers' quality of the interpretation of students' work and practice of FA in general improved (Gearhart et al., 2006; Herman & Choi, 2008; Ruiz-Primo & Furtak, 2006, 2007). Effective formative assessment, therefore, requires that students and teachers diagnose students' current understanding and compare that to the learning goal and follow a clear path to meet this goal.

Status of three-dimensional teaching, learning, and formative assessing

Research on three-dimensional teaching and learning is still in its infancy, as reflected by the paucity of studies in this area. There is uncertainty as to how the three dimensions are to be packaged into classroom instruction, given that the reform standards do not equate the curriculum (NGG Lead States, 2013). Teachers will have to shift their thinking to embrace the integrated framework to guide how they plan and implement three-dimensional instruction. This shift, according to NASEM (2017), requires weaving the three dimensions together in all aspects of science so students can understand how science works and continue to use the knowledge throughout their personal and professional lives. The climate in the classroom will change from helping students "absorb sets of factual knowledge to strengthening students' capacity to think and reason about ideas and information they are tackling" (NASEM, p, 14). Integral to three-dimensional instruction is a new way of formatively assessment that allows the teacher to track students' progress. The teacher should turn their attention from assessing isolated ideas to assessing how students use SEP in the context of CCC and DCI with a variety of challenging tasks, so students can focus on growing their abilities and perceptions into an incomplete understanding of the standard (NASEM, 2017).

Studies in this review focused on instructional resources were helpful to both teachers and students in improving their understanding of the dimensions. However, teachers are unaccustomed to these new approaches and struggled with the new materials and how to align their lessons to NGSS. Teachers and teacher educators will need directions in the form of engaging resources and student-driven material. Studies that provided professional development to pre-and in-service teachers indicated a positive impact on teachers' understanding of the three-dimensions. However, continuous and specific training and practice are necessary to build teachers' knowledge and shift their focus towards helping students figure out how and why phenomena occur and construct solutions to problems. More studies are needed on professional development strategies of 3D learning to provide maximum benefit to teachers. The recommendation for assessing 3D learning included using multicomponent tasks that measure complex reasoning. Studies on exploring assessment strategies, especially the formative assessment of 3D learning, must accompany efforts to create integrated instructional materials.

Limitations and Strengths of the Studies

A glaring limitation in most of the studies reviewed is the small samples, making the studies non-representative and the results non-generalizable. Anderson et al. (2007), Buck et al. (2009), Herman et al. (2010), Lee and Wiliam (2009), Offerdahl and Tomanet (2011), Ruiz-Primo and Furtak (2007), and Sato et al. (2008) all elected small sample sizes to study their variable in detail, the result is non-generalizable because each teacher’s context is different. However, the small sample size is also a strength because it allows the researcher to obtain a detail account of the case. Another limitation observed was the use of measures which were in some cases not valid (Hickey et al. 2012; Sato et al., 2008; Wiliam et al., 2004), imperfect (Herman et al., 2006; Herman & Choi, 2008), different measures used for the same variable

(Hickey et al., 2012; Furtak & Ruiz-Primo, 2008) and some measures were not pre-tested (Jenkins, 2009; Penuel et al., 2007). Furthermore, the design of some interventions involved unique context and pilot instruments, making them less transferable to another context (Furtak &Heredia, 2014; Gearhart et al., 20006; Feldman & Capobianco, 2008; Morrissette, 2011; Offerdahl & Tomanek, 2011; Ruiz-Primo & Furtak, 2006). Another strength in some of the studies lie in their unique design, that can serve as methodological guide for future studies. For example, Ruiz-Primo and Furtak (2007) ERSU model can be used for informal assessment of classroom conversations.

Directions for Future Research

Reflecting on the studies, a noticeable gap is the absence of studies on students' self-assessment during FA. It may be an omission, but the search yielded no articles that discussed students' self-assessment with FA. Because the students play a significant part in their learning, the capacity to monitor their learning and act on the feedback provided by the teachers is a necessity for assessment for learning. Therefore, studies to document better ways to foster students' self-assessment practice in taking ownership of their learning is vital. The articles selected did not also look at teachers' expertise in implementing self-created FA tasks. Yin et al. (2008) suggest that when teachers are partly or wholly responsible for assessment design that is tailored to their needs and preferences, they will feel comfortable using their assessments.

Given that most of the studies reported teachers limited conceptual understanding and practice of FA (Buck et al., 2010; Herman & Choie, 2008; Yin et al., 2008), there is undoubtedly a need for future research in this regard. Another gap that needs attention is further research to refine the many variables that have been used to investigate FA understanding and its practices. Another study that demands the attention of every school district is to create professional development interventions alongside action research for science teachers to receive specific training and practice the knowledge and strategies in their classroom while simultaneously reflecting and documenting their experiences.

Suggested Questions to Pursue

The questions listed below arose from the recommendation of the articles and from identifying what was left out or inferred from their conclusions and implications.

- What is the picture of FA in today's science classrooms, and how is teacher understanding reflected in practice?
- How do teachers create classroom culture to improve dialectic discuss among students and their self-assessment capacity?
- What are teachers' experiences with three-dimensional teaching and formative assessment?
- What impact would an assessment created on the cognitive demand of the task, student ability, task efficiency, or task fitness have on student achievement?
- What are some effective strategies inherent in a professional development program to entice science teachers to embrace FA practices?
- How will supporting the process of collecting and using assessment results impact teachers' practices of informal FA and student participation?
- What is teacher experience with the implementation of self-designed FA tasks, and can this authorship encourage continuous tinkering?
- What are high school science students' understanding of self-assessment and the academic discourse upon which the language of formative feedback is based?

- How are science teachers navigating the new terrain of three-dimensional teaching with respect to FA?
- What is the product of learning associated with the development of integrated assessment knowledge?
- How can students apply their understanding of CCC from one content to another or make a connection between ideas?
- What type of support is more effective for the teacher to grapple with complex scientific practices?

3. METHODOLOGY

Introduction

> Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited (Black & Wiliam, 2009, p. 7)

The essence of Black and Wiliam's argument is that one can assess practice in the classroom as formative based on whether all interacting parties collect and analyze data about student performance and use the results to provide a better teaching and learning opportunities. An understanding of what it means for assessment to be formative gives a starting point of what to look for in such practices. The purpose of this study is to explore science teachers' understanding of formative assessments and how they elicit and interpret students' integrated science knowledge to adjust instruction and improve students' three-dimensional learning. The study will also explore teachers' understanding of three-dimensional teaching and the strategies and/or obstacles involved in implementing the formative assessment. Constructionism will serve as the theoretical framework and situated learning as the conceptual/disciplinary framework for this study.

Theoretical Framework

Formative assessment is necessary to foster reflective teaching and improve students' learning (Black & Wiliam, 1998; Furtak & Heredia, 2014). A theory of formative assessment requires a framework that clarifies the process and allows for monitoring and improving learning and instruction (Taras, 2010). Furthermore, a theory of formative assessment can provide the epistemological structure for the subject matter, teacher's professional knowledge, a frame for teaching and learning, and the theory of communication (Yorke, 2003). A learning theory consistent with teachers helping students operate within a framework of formative assessment must be guided by three questions. Where is the student going? Where is the student now? How to help the student get there? (Black & Wiliam, 1998, 2009; Sadler, 1989). However, formative assessment does not ascribe to any given theory of learning (Black & Wiliam, 2009; Sadler, 1989) and interactions in the classroom (Black et al., 2003). This elusiveness can be attributed to the use of multiple dimensions to assess students' performances, use of improvised instruments with no theoretical bases, and lack of teachers' theoretical grounding and inadequate knowledge of assessment practice (Yorke, 2003).

Some researchers used a formal theory to understand the practice of formative assessment; others used informal or personal framework for the same cause. For example, Lee and Wiliam (2004) and Olferdahl and Tomanek (2011) used a constructivists theory to describe teachers thinking/construction of knowledge of formative assessment. Similarly, Anderson et al. (2007) and Furtak and Heredia (2014) used a situated lens to explain teachers' understanding of formative assessment based on participation and context, respectively. Conversely, some researchers develop their assessment framework to describe the facets of formative assessment being observed. Furtak and Ruiz-Primo (2008) developed a theory of formative assessment prompts, Ruiz-Primo and Furtak (2006, 2007) used a discourse model, and Herman and Choi (2008) used a quality goal model. These studies all have in common a vehicle for eliciting students' ideas, interpreting the ideas, and acting on the evidence. In this study, a social constructionist theory will be employed as the theoretical framework and situated learning as a conceptual or disciplinary framework.

Social constructionism

Social Constructionism is an epistemological view of how knowledge is constructed and understood in a social context. It examines how “meanings are constructed by human beings as they engage with the world they are interpreting” (Crotty, 1988, p. 43). Social constructionism is an inquiry process, a system to understand the world by experiencing it in social practices. The emphasis is on the interaction between individuals and the construction of reality based on contextual and linguistic/discursive factors. According to Merriam (2009), “individuals construct reality in interaction with their social worlds” (p. 22). Individual learning thus occurs within the interaction and best with meaningful activity involving the learner in constructing a tangible product. The interactions from which data is generated also enhances the analysis. This theory assumes that teachers can enhance their knowledge of practicing and assessing three-dimensional teaching by collaborating and or practicing in their classrooms. With a focus on relations, teachers can construct mental models based on experience and perception to understand the formative assessment of 3D. Teachers can collaborate with each other to develop and implement shared functional meaning such that realities are socially negotiated through their collective experiences and interactions (Andrew, 2012; Raskin, 2002).

Social constructionism permits flexibility in the research design. The researcher paid attention to the voices of the participants in the study and the evidence they shared about their practices upon which they build the meaning of 3D and formative assessment. This lens helped to describe teachers’ understanding of concepts in the social context of their classroom, and the “researcher is interested in understanding the meaning a phenomenon has for those involved” (Merriam, 2009, p. 22). How are FA and 3D teaching understood and practiced from the experiences of the teachers? This contextualization permitted the recognition of multiple realities and multiple interpretations of the concepts of three-dimensional teaching and learning used in the collection and analysis of data in this study. Crotty (1988) acknowledged that “there are no true or valid interpretations. There are useful interpretations for sure, and they stand over against interpretations that appear to serve no useful purpose” (p. 47). This lens guided the exercise of useful interpretations of the concepts in this study. The relationship between participants and the researcher was that of mutual respect. The teachers were experts in their practices and were autonomous and reflective in all their decisions. The researcher was transparent during data collection and analysis. The researcher interpreted the data, thus influenced and affected the research in the process of knowledge precaution and acknowledged that objectivity is not feasible nor desired. The researcher assumed the position of not knowing, which promoted curiosity with the investigation and was willing to accept data that were or not fitting to prior knowledge or their own experience. This theory was thus apt to provide a lens through which the researcher defined and categorized teachers’ experiences from the perspective of the teachers.

Situated learning

The situated view is an open-minded theory that highlights the importance of context and lived experience or our thoughts and actions (Dewey, 1998). This view illuminates a shift from individual behavior and cognition to a larger activity system that coordinates all components together, the teacher’s informational structure and science practices, the learner, and tools in the context of the activity (Greeno, 2006; Putnam & Borko, 2000). Situated view stipulates that what is learned depends on how it is taught or on how the learner interacts with the information in the context in which learning occurs (Gee, 2004). Participation requires active negotiation and construction of cognitive performance thus developing acceptable knowledge. Accordingly, learning and knowing are situated, socially embedded, and distributed (Greeno, 2006; Putnam &

Borko, 2000). Learning in the context of authentic activity is unintentional (Lave, 1988). Learners use their prior knowledge in a specific subject and apply the knowledge to a similar context (Lave & Wenger, 1991).

Situated learning provided the lens on how to conceptualize the exploration of three-dimensional teaching and learning. It guided the link between the research questions and the appropriate method. Relevant to this study was the assumption that teacher's learning occurs in context, is situated, and is achieved in and with interaction. It was also assumed that teachers would apply their prior instructional skills to 3D teaching and formative assessing. The teacher's formative task developed, phenomena used, or communicative practices used influenced the way three-dimensional teaching was enacted and assessed. According to Lave (1988), "activity in which knowledge is developed and deployed … is an integral part of what is learned" (p. 32). Meaning making from the authentic activity occurring in the classroom community is through connecting prior knowledge to a new context. The activity, concept, and culture are interdependent. When the teacher presents students with familiar phenomena, the students can gradually build their knowledge of the world in which the concepts are used and of the concepts themselves.

Conceptual knowledge, according to Brown, Collins, and Duguid (1989), is like a set of tools. They suggested that "to learn to use the tools of discipline as practitioners use them, a student, like an apprentice, must enter that community and its culture… learning thus is a process of enculturation" (p. 33). The teacher involved students in a kind of learning that included integrating the three dimensions to move students from a novice toward an expert stance along a continuum. The student thus could reason using everyday models, not laws; act with conceptual situations, not symbols; solve ill-defined problems, not well-defined problems; produce negotiated meaning, and socially construct understanding, not fixed meaning and immutable concepts (Brown et al., 1989). The result is a shift from acquiring inert knowledge to figuring out and developing useful and robust knowledge. Assuming that concepts/practices are situated in activity and increasingly developed through participation, the teacher and student knowledge of concepts "continually evolve with each new occasion of use, because new situations, negotiations, and new activities inevitably recast it in a new, more densely textured form" (Brown et al., p.33). Knowing and doing thus leads to original solutions. The situated view is important for its ability to use multiple frameworks or lenses to govern the choice of data collection and analysis tools and to place emphasis on context-based questions and resources available (Borko, 2004). This study used multiple lenses based on a domain and a participatory context (Anderson et al., 2007). The domain context used a near vision based on situatedness (teacher cognition linked to the environment) to collect data on teachers' understanding of three-dimensional teaching and learning and views of formatively assessing these practices. The participatory context used a far vision based on a constructionist's pedagogy to collect data on social interaction in the context of each teacher's classroom culture.

Conceptual Framework

When an assessment task is developed with a learning goal in mind, it elicits student's ideas at the level of success criteria and provides the teacher with the evidence necessary to adjust instruction for deeper learning. Formative assessment solicits active student involvement as crucial to its success; students, therefore, can develop evaluative expertise for self and peers (Sadler, 1989). Providing students with corrective feedback ensures that students are willing and able to build on the directives to 'feedforward' (Carless, 2007). The processes of identifying/eliciting, collecting, and interpreting/analyzing data to evaluate a student's level of performance regarding core ideas with three-dimensional teaching and learning require that the teacher employs direct, indirect, quantitative, and qualitative measures necessary to meet the learning objective. It is, therefore, possible for teachers to use formative assessment core elements to facilitate three-dimensional instruction. To measure students developing understanding of 3D learning, assessment tasks must 1) examine how students use SEP in the context of CCC and DCI, 2) use many challenging tasks to provide multiple opportunities to demonstrate learning, 3) elicit diverse and specific information for next step of instruction by the teacher and for student to monitor their progress, and 4) focus on students' conceptual development rather than on right or wrong answer (NAP, 2017).

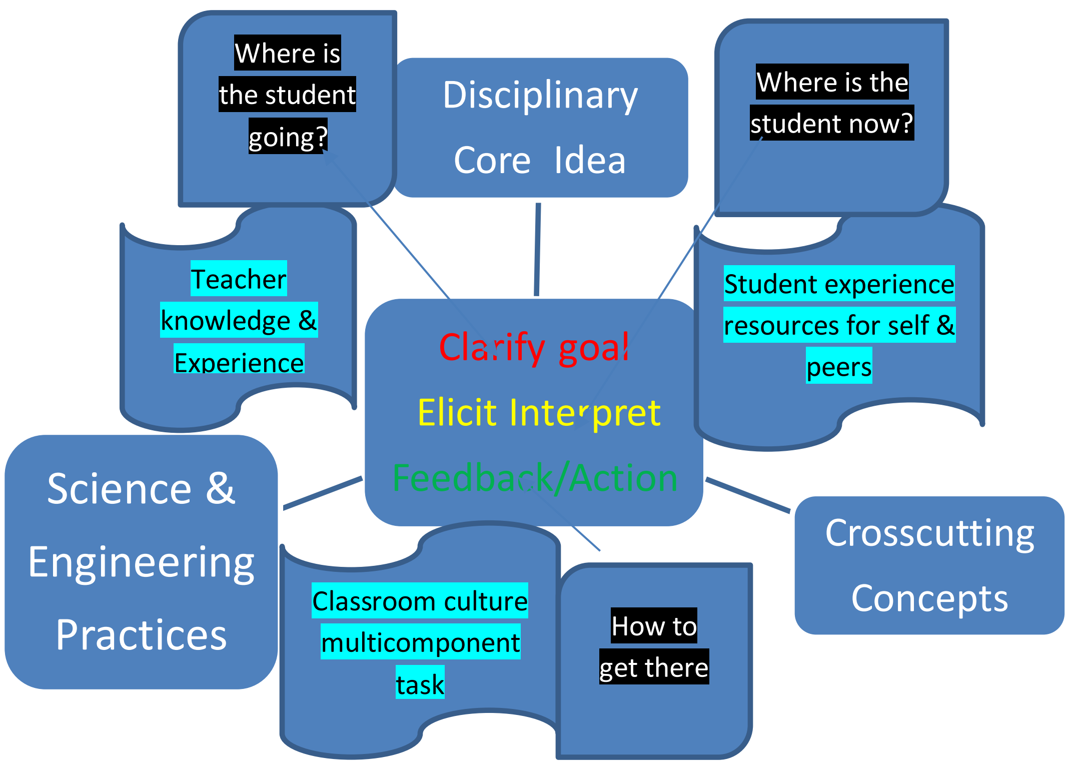

Figure 1. Integrating practices with assessment elements

Figure 1. above proposes a framework for formative assessment of three-dimensional teaching and learning. It entails using a multi-component task to assess students understanding of

phenomena through integrating DCI, SEP, and CCC. The task clarifies the big idea focused on how and why phenomena (scenarios of natural events) occur. Task elicits thinking and scaffolds students to develop and build habits of mind for scientific understanding and uses crosscutting themes to make connections between the SEP and DCI and links concepts across disciplines. This process is at the mercy of a teacher's knowledge and experience and ability to create a classroom culture of figuring out or identifying problems, asking questions, constructing solutions, and that encourages students to take ownership of their learning and be resources for each other. Integrating the three dimensions would provide the teacher with a means to elicit a complete and accurate explanation of the phenomenon using CCC as a lens to interpret student's thinking and provide the next steps for learning.

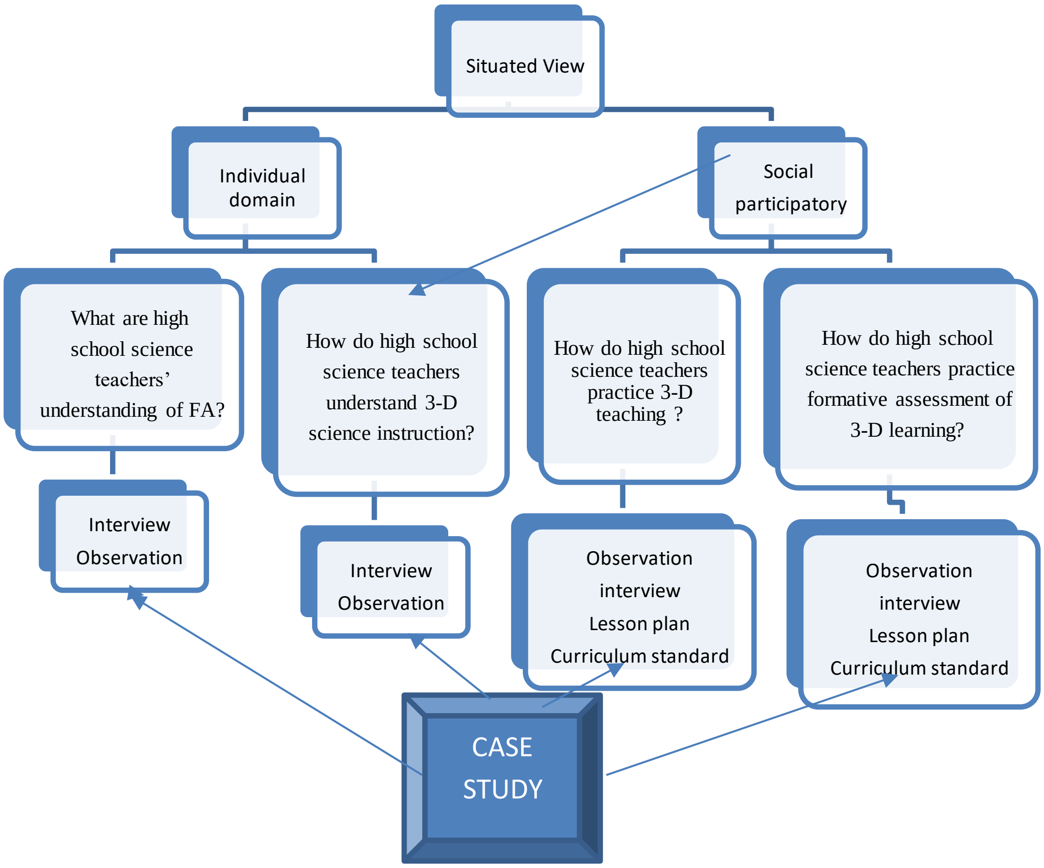

Figure 2. Methodological Framework for exploring FA of 3D Teaching/ Learning.

The situated lens guided the collection of data on formative assessment and 3D teaching in the context of both the teacher and classroom interaction. Figure. 2 proposes how the situated lens steered the researcher through the process of linking the research questions to the data collection methods, to explore the concepts of formative assessment for three-dimensional teaching and learning. The situated perspective afforded the researcher with the lens to identify specific elements as evidence of the teacher's personal and social construction of the concept of three-dimensional teaching and learning. It also directed the researcher's attention towards the teachers' formative assessment tasks and the formative interactions occurring in the classroom.

Qualitative Study Methodology

Research, according to Merriam (2009), is "inquiring into or investigating something in a systematic manner… to know more about something than we did before engaging in the process" (p. 3). Research can be quantitative or qualitative. Qualitative research is a "situated activity that locates the observer in the world" (Denzin & Lincoln, 2005, p. 3). Qualitative research is grounded on four characteristics, "the focus is on process, understanding, meaning; the researcher is the primary instrument of data collection and analysis; the process is inductive; and the product is richly descriptive" (Merriam, 2009, p. 14). A qualitative researcher studies "things in their natural settings, attempting to make sense of, or interpret, phenomena in terms of the meaning people bring to them" (Denzin & Lincoln, 2005, p. 3). Merriam divides qualitative research into two categories, basic and applied based on its purpose. Within these processes, basic research is "motivated by an intellectual interest in a phenomenon and has as its goal the extension of knowledge… Applied research is undertaken to improve the quality of the practice of a particular discipline." The basic qualitative researcher is attentive to "(1) how people interpret their experiences, (2) how they construct their worlds, and (3) what meaning they attribute to their experiences" (Merriam, 2009, P. 23). This study employed basic qualitative research to know more about teachers' understanding and practice of FA and 3D teaching, and the study "may eventually inform practices." (Merriam, 2009, p. 3). Given that the unit of analysis in this study is a bounded system or a case, the study methodology will be a qualitative case study.

Case Study Design

A qualitative case study can be defined by the process of investigation, by the unit of analysis or by the product of the investigation. Case study as a research method, is defined in terms of its scope and features as, “an empirical inquiry that investigates a contemporary phenomenon (the case) in depth and within its real-world context, especially when the boundaries between phenomenon and context may not be clearly evident” (Yin, 2014, p. 16). Some definitions lean more towards the unit of analysis. Stake (2005) suggests that a case study be defined by the “interest in an individual case… a choice of what is to be studied” (p. 443). Case study can be presented in varying forms. Stake (2005) differentiates three types of case study, instrumental- the phenomenon is examined to provide insight on a different issue or make generalization from; Intrinsic- pursued based on researcher’s intrinsic interest on a case; Collective- uses multiple case studies to investigate a phenomenon. This study employed a multiple case study to investigate formative assessment and three-dimensional teaching. The participants in this research method, “construct reality in the interaction with their social worlds” and the researcher was “interested in understanding the meaning a phenomenon has for those involved” (Merriam, 2009, p. 22). The researcher explored teachers understanding of three-dimensional teaching and how they created and implemented formative assessment in their daily practices. This method was attentive to how the researcher’s subjectivity could influence the results, and how the process allowed for iterative collection and analysis of rich data (Merriam, 2009).

With qualitative case study method, the researcher strived to understand secondary science teachers’ knowledge of formative assessment or the meaning they attribute to everyday practice of formative assessment with respect to three-dimensional learning. A case study methodology can show casual relationships between teacher-created formative assessment tasks and their instructional practices (Yin, 2013). Case study focuses on contemporary problems in context such as the bounded classroom to yield deep understanding of the phenomena through an in-depth analysis of the system (Merriam, 2009). For example, exploring individual teachers' classroom practice of formative assessment fenced within three-dimensional teaching and learning could illuminate teachers' knowledge of formative assessment and its practice within a reform-based context. The richness of data obtained in a case study may also generate new thinking and new ideas related to formative assessment and three-dimensional teaching. This study employed a variety of data sources to establish a chain of evidence and specific approaches to data analysis for a robust finding (Yin, 2014). Data collected from the different cases allowed the researcher to explore varieties within and between cases such that findings are replicated across cases (Baster & Jack, 2008) to provide compelling evidence resulting in a more robust study. Besides, the different teachers had different contexts provided enough rich and detailed data for each to be a standalone case to reveal corroborating information (Yin, 2014).

Although case study is favored for its use of replicative logic, data in this study were collected simultaneously for all teachers to align with constructionists' views, and this study profited from parameters identified in a pilot study for guidance. This process provided flexibility to facilitate and document the understanding of the research questions and generated new thinking and insights related to formative assessment. The method thus helped to describe patterns and relationships that emerged from teacher understanding and practice; it also helped to identify questions and outline how to answer them or in other words, to frame the whole study. A case is a unique situation or occurrence able to provide rich information about an area of inquiry or the activities and experiences of those involved and its context (Stake, 2000). The case in this study is high school science teachers and their understanding and practice of formative assessment in the context of three-dimensional teaching. This chapter provides information on the specific direction to explore the research problem, describe the relationship between variables, identify and outline the development of the study. The goal was to gather enough data to capture the teacher's knowledge and enactment of formative assessment with three-dimensional teaching. To accomplish this task, the researcher borrows an insight from past studies to design how to answer the following questions:

1. What are high school science teachers' understanding of formative assessment?
2. How do high school science teachers understand three-dimensional science instruction?
3. How do high school science teachers practice three-dimensional teaching?
4. How do high school science teachers practice formative assessment of three-dimensional learning?

Participants

This study took place in a large suburban district in the Southeastern United States with a majority population of students from underrepresented groups. All the schools in the district were title one schools (serving low income), plus all the students were entitled to a free breakfast and lunch. A purposive sample (Guba & Lincoln, 1994; Yin, 2014) was selected because the teachers could provide, or from which one could learn a great deal with respect to the intent of this study. Another criterion for selection was based on respondents and the location of the schools. Three secondary science teachers were solicited from two neighboring high school. One teacher, Andria, from the first school (on the East High) was referred by another teacher in a pilot study as a "good teacher." Two teachers were selected from the second school (on the West High), the first Chelsey, was selected for previously attending professional development on 3D teaching. The second teacher, Paul (department chair) was referred by Chelsey because "he is into it" (the 3D teaching). All three teachers are currently teaching biology which has an end of course test (Georgia Milestone Assessment Test (GMAS)). These teachers had each attended two county-wide professional development sessions provided to all science teachers in preparation for adopting 3D teaching in addition to two school-level in-service training exclusive for GMAS teachers. This non-probability group was chosen with the assumption that it is a logical representation of the population and will provide the necessary data. The choice of high school science teachers is because they are the population of interest for this study. Additionally, teachers with similar content and context will make it more likely to construct the case, to focus on teachers rather than groups, to replicate the procedure with high flexibility, and to limit the number of factors that could influence the outcome of the study (Levy, 2008). Teachers were also chosen using a query based on years of teaching. One teacher with five years of teaching experience and the remaining two with 10th – 20th years of teaching experience. Andria was chosen with the assumption that her pre-service training curriculum had assessment as a core part of its coursework. Chelsey and Paul were chosen with the assumption that with or without coursework, they were experienced teachers and have had some professional development and practice on formative assessment in their classrooms. Veteran teachers (with 21 years and up) were excluded as participants. The assumption was that most of them attended teacher training programs when assessment was not offered as a course of study, even though FA might had been recognized as a tool to enhance instruction and students' learning (Darling-Hammond & Bransford, 2005). In addition, this group of teachers may have long established a way of teaching dictated by their apprenticeship of observation (Lortie, 1975) and pressure from high stakes testing with the need to cover the curriculum (Yin et al., 2008). The rationale for this purposeful sampling is to increase the chances that the participating teachers possess the knowledge and skills of formative assessment that will provide the opportunity to collect the necessary data to answer the research questions (Gearhart et al., 2006). Teachers selected met the expectation to yield similar data. The choice of three teachers was suited to extract rich data to help understand the complex interrelationship within the context (Yin, 2013). Access to the schools was gained through IRB approval from the county and site approval from the respective school principals. Additional assistance was requested from the department chairs of each school. Table 1. below summarizes the characteristics of the participants, their years of teaching, ethnicity, degree and major, method of certification, and the current course taught.

Table 1.
Participants and their descriptions

Teacher	Years of experience	Ethnicity	Degree Major	Major Certification	Teaching position
Andria	5	Black Female	Masters Biology	Traditional route	Biology
Paul	16	White Male	Specialist Biology	Alternate route	Biology, regular & AP
Chelsey	14	Black Female	Specialist Broad field science	Traditional route	Biology Environmental, regular

Data Sources

Data in qualitative research is usually conveyed through words. This data can be in the form of direct quotations from interviews, a description of the participants' activities, behaviors, or actions from observations, and substantial from documents (Merriam, 2009). The first source of data was interviewing. An interview is a systematic engagement in conversation activity between the interviewer/researcher and interviewee/participant guided by focus questions. The study conducted many informal conversations that were open and adaptable to the interviewee's ways and priorities and two semi-structured interviews. The questions for the semi-structured interviews were gathered from the informal interviews, observations, and from the literature. These interview questions were less structured and presented in a non-threatening or in a friendly manner. They were worded to guide the interviewee and to make the researcher “appear genuinely naïve about the topic and allow the interviewee to provide a fresh commentary” (Yin, 2014, p. 111) on the query. A semi-structured interview allowed for the flexibility of participants’ responses and the opportunity to capture diverse perspectives.

The second source of data collection was observations- a systematic inquiry addressing a specific question and was subject to checks and balances from the observer to yield a trustworthy result. Good qualitative research observation demands selective attention to a few crucial things from a researcher that might escape the attention of others (Merriam, 2009). Assuming the position of observer participant, my activities as an observer were made known to the participants and took precedence over my role as a participant. Eight observations each were conducted per case to capture the teacher’s practice of formative assessment and 3D teaching.

The third source of data collected was documented (lesson plans, performance activities, assessment tasks, curriculum standards, and field notes). These documents are “readymade source of data easily accessible to the imaginative and resourceful investigator” (Merriam, 2009, p. 139). Documents were included because information can be learned from it to stimulate inquiry through observation and interviews. Documents can be a good source of data because it “does not intrude upon or alter the setting in any form. Nor are documents dependent upon the whims of human beings whose corporation is essential for collecting good data” (p.139) compared to interviews and observations. The documents were examined to improve the understanding of teachers’ thinking and preferences of practice. The fourth data source was audiotapes of the classroom conversations to capture the interactions and type of discourses occurring in the classroom.

Procedures

Data collection

The process for collecting data occurred in and out of the classroom. Data were obtained using interviews, observations, and documents (lesson plans, assessment tasks, and written notes) as data collection tools (Guba & Lincoln, 1994). The interviews were conducted in each teacher's classroom at a convenient time during their planning period and after school. The observations, however, occurred during school hours when the teacher was in the classroom teaching.

The study started with an informal interview to capture the teachers' beliefs about teaching and the goal for student learning. Following the open talk, observations were conducted twice a week per teacher, a total of eleven observations each. The intervals provided the researcher time to review and reflect on what was observed and generate a further point of inquiry to inform the next observation and subsequent interview (Penuel et al., 2007). The researcher made sure to observe the beginning, middle, and end of the curriculum unit to capture a realistic picture of teachers' practice. The researcher took notes and audiotaped the observations. The duration of each observation was 55-minute, the length of a high school class period, to capture the beginning and end of the lesson. Additional informal interviews were conducted before or after observations or when possible. Documents (teachers' lesson plans and assessment tasks) were collected after each observation or when available, physically and/or through email, to gather more information on the practice. The observation field notes and documents served as a window into the teachers' practices, and the researcher sought clarification from teachers when necessary. Information from these documents provided specific details to corroborate with other data sources (interviews, audiotapes, and observations data) to make inferences about the findings. The first semi-structured interviews were conducted after the fourth observations and the second at the end of the observations, as a guided conversation with a fluid, unstructured line of inquiry from open-ended questions. The semi-structured interviews ranged from 11 to 50-minute-long and audiotaped to capture every statement. The duration of this study was a semester to allow time for the researcher to collect and analyze interim data on a unit and follow up data using insight from the draft. Table 2. below summarizes the data collection sources and procedures.

Table 2.
Data collection procedure and duration

Stages of data collection	Length (minutes)	Action taken	Context
1. Informal interview	Varies	Open conversation directed by participant interest and priorities including their role as teacher, goal for students, decision on what to teach, and when to move on.	Non-instructional time at location of participant's convenience
2a. Observations, eight to ten per teacher, (once to twice a week for 6 weeks)	55	Document what participants say and do. Collect related documents for more evidence	Instructional time Classroom setting
2b. Collect teacher's lesson plans and assessment task	Varies	Researcher analyzes documents for additional information and to clarify what is heard or observed	Non-instructional time, teacher determines when and how
2c. Audiotapes of lesson and interviews	55 11-35	To capture classroom conversations as corroboration of what researcher hears and see	Instructional time Classroom setting
3. Researcher reflective field notes		Jot down pertinent information during observations, After interviews and reading documents	When occasion arises
4. Semi-structure interview	11 to 37	Guided but fluid conversation about knowledge and practice of formative assessment and new standard. Participant share anything about the new curriculum, from role, assessment, etc. Questions will come from the open interview, observation, and the literature to guide interview (see appendix)	Non-instructional time at location of participant's convenience

Data analysis

Data analysis serves to make meaning of anything that a researcher gathers to answer the research question(s). According to Merriam (2009), it is a "complex process that involves moving back and forth between concrete bits of data and abstract concepts, between inductive and deductive reasoning, between description and interpretation" (p. 176). The instrument in qualitative research is the researcher, who uses his/her filter/lens to interpret the findings (Patton, 1990; Yin, 2014). As such, the results will originate from personal interpretation of evidence from data shaped through experience leading to multiple perspectives. The purpose of this analysis was to make sense of the data by consolidating, reducing, and interpreting (Merriam, 2009; Offerdhale & Tomanek, 2007) what the teachers communicated and what the researcher gathered. A quality analysis requires that one maintains both an open mind and an interpretive and reflective disposition. Case study as an iterative process of data collection and analysis (Anderson et al., 2008), allows for the simultaneous collection and interpretive analysis of data (Merriam, 2009; Offerdhale & Tomanek, 2007). According to Merriam (2009), "Analysis begins with the first interview, the first observation, the first document read. Emerging insights, hunches, and tentative hypothesis direct the next phase of data collections, which in turn leads to the refinement or reformulation of questions" (p. 165). This analysis was, therefore, critical to convey an understanding of the case. Table 3. below highlights how the interview data were analyzed.

Table 3.
Interview data analysis procedure

	Open code Relevant words, phrases,	Axial codes Groups/theme	Thematic analysis Reviewing, interpreting
Formative assessment	Check, know where, need to go next, quick, short interval,	Clarifying Eliciting Interpreting Using feedback	Identify and share success criteria Use tasks to elicit and reinforce learning Actions for next steps
Three-Dimensional teaching	Integrate all three, phenomena, relatable, interest, what scientist do, figure out,	Engaging in discussion Involving in the practices	Merging the practices of each dimension to make connections, apply knowledge, figure the natural world
		Repeat, regroup	Repeated reviewing and interpretations

The unit of analysis or the case in this study is the teachers (Lee & Wiliam, 2005), and case summaries were developed for each. After data collection, all the materials were organized into a case study database (Yin, 2014). All the interview transcripts were arranged under each teacher in a word document, and all the observation transcripts were also organized per teacher in a separate word document. These data were stored in the researcher's computer and password protected. The interview data were reviewed and compiled into a descriptive narrative (Penuel et al., 2007) and case summaries (Furtak & Heredia, 2014). Interview data were transcribed verbatim, and anything that might seem important was noted. This open coding was used to assign comments to relevant concepts by identifying and developing categories with an open mind (Merriam, 2009). These categories were then assembled into themes (Charmaz, 2006) of formative assessment and three-dimensional teaching elements. The themes were assigned descriptive codes (Gearhart et al., 2006; Herman & Choi, 2008) that represent evidence of formative assessment understanding based on Wiliam (2010) five key FA strategies and the Framework for K-12 Science Education (NRC, 212) three characteristics of developing assessment tasks. Similar concepts were grouped together as axial coding, conducted to analyze how these themes relate to the concept of formative assessment being investigated. The process of linking the data to themes was through pattern matching of what emerged from the data, building explanation from the different responses obtained, and a cross-case synthesis of the different cases (Furtak & Heredia, 2014, Lee & Wiliam, 2005). Claims and generalizations across cases were evaluated for accuracy, and its content dissected to obtain accurate and clear meaning (Merriam, 2009).

Classroom conversation audio transcripts were analyzed for evidence of formative assessment elements and to reveal the different discourse patterns and underlying meanings hidden in classroom talk (Lemke, 1990). Segments and vignettes from these transcripts were marked to provide support for themes identified in the interview transcripts. The researcher also analyzed the classroom audio transcripts to understand teachers' informal assessment practices, how or whether the teachers clarified the learning goals, elicited evidence, provided feedback to students, or activated learners as resources for each other. Also, the types of interactions the teachers initiated with the students were identified and categorized into either broken or complete cycles to understand how classroom discourse was used.

Trustworthiness

This study established a transparent chain of evidence so that other researchers can reconstruct the steps from questions to the conclusion and used participants and expert checking to improve construct validity (Furtak & Heredia, 2014) or to legitimize inferences. Data were collected through different sources including observation, interviews, and documents; using different strategies, such as opened unstructured interviews and observations before interviewing; triangulating data from different sources by checking for consistency or disconformity of findings to obtain multiple evidence (Offerdhale & Tomanek, 2007; Patton, 2002).

Credibility for this study was by cross-verifying theory or claims from different teachers (Furtak & Heredia, 2014; Penuel et al., 2007) with multiple perspectives (Herman & Choie, 2008) to gain a better understanding of different dimensions of formative assessment. The interview data were collected and transcribed verbatim to make sure the analysis is based on valid information and to avoid infusing foreign ideas into the study. The interview process, according to Roulston (2010), should be reflexive, researchers are "self-consciously aware of their subjectivities in relation to the research participants and the research topic" (p. 89).

Participants verified transcribed and analyzed interviews for accuracy, so themes and thoughts were correctly reported (Furtak & Heredia, 2014; Offerdhale & Tomanek, 2007). Observation in this study was unobstructed, conducted with a trained eye to focus on relevant occurrences, and the researcher kept detailed notes. Repeated observations, along with data analysis and reflections, were conducted. Multiple interpretations arose from different contextual perspectives of findings to minimize bias and relate theoretical idea(s) that may develop to previous studies and enhance credibly.

The researcher established confirmability by maintaining an audit trail and was reflexive by questioning her thoughts and actions throughout the study. A detailed account of data collection, analysis, and interpretation was provided, and the researcher recognized how one's presence and perspectives might unknowingly influence participants' responses and behavior. Dependability for the study was achieved by providing detailed accounts of data collection procedures, using a case study protocol, revealing theoretical assumptions to make explicit the process, and developing a case study database to guarantee that the study could be replicated. The repeated analysis was used for consensus of interpretation of data to minimize errors and biases (Furtak & Heredia, 2014). The method and procedure for data collection were consistent for all participants to assure reliability and the results compared with previous findings to ensure an inquiry audit (Bryman, 2008). However, although planned to be similar, due to social constructionism, no two cases were the same. Efforts to achieve transferability for the study included taking a nested approach by using all high school science teachers, providing a rationale for selecting these teachers, providing enough detail on the case study context, and using the same variables with all participants.

Ethical procedure

The participants were treated with the utmost respect, that is, by considering each teacher as the master of their classroom. Consent was secured through IRB approval for all activities involving the participants. Confidentiality was achieved through the cautious handling of data by assigning pseudonyms to each teacher. The data were stored in a secure place, and personal information was kept private.

4. RESULTS

The purpose of this study was to explore science teachers' understanding of formative assessments and how they elicited and interpreted students' integrated science knowledge to adjust instruction and improve students' three-dimensional learning. The study also described teachers' understanding and practice of three-dimensional teaching from the perspectives of the teachers. This research was grounded on four exploratory questions to guide the trajectory of the study consciously.

1. What are high school science teachers' understandings of formative assessment?
2. How do high school science teachers understand three-dimensional science instruction?
3. How do high school science teachers practice three-dimensional teaching?
4. How do high school science teachers practice formative assessment of three-dimensional learning?

Three teachers were involved in this study. They were willing to give their perspective of the concepts of formative assessment and three-dimensional teaching and learning and to open the doors into their classrooms to reveal how they enacted these concepts. Chelsey, Paul, and Andria (pseudonyms) were interviewed, observed, and documents collected from each of them. Although the idea of FA can be unenticing to some teachers and 3D teaching may be considered as innovative reform, the teachers in this study were confident in their knowledge and flexible and adaptive in teaching science, to be willing to share their experiences. The narrative below provides a synopsis of these three teachers' understandings and their practices. Data from interviews were identified with the letter (I), that from observation with (O) and from documents with (D).

Case 1. Chelsey

Chelsey is in her 15th year of teaching and has taught various grades from 6th up to undergraduate. She obtained her teaching certification the traditional route by attending a teacher education program. She is certified in broad field science and is working towards her doctorate. Chelsey is currently teaching ninth grade Biology and Environmental science and has been teaching Biology for the last five years.

Teacher's Understanding of Formative Assessment- Checkpoints

Chelsey described FA as using a variety of spotters like "raise your hand, thumbs up/down… green light, red light. The red light means oh hold on I don't understand I need help and the green light means I get it, go on. Also ticket out the door, high five, and short answers" (I). Chelsey considered all these as "formative" because it helped her know what the student had learned. These examples of formative assessment tools indicated that Chelsey understood FA as something that provided quick Information about the status of student learning. She further explained that it is easy to do the raise your hand or thumbs up/down type spotters because "it is part of their body and does not need preparation to use" (I).

Chelsey believed that FA is a process that helps show "where our students are and to gauge their understanding of the information taught, to get some background knowledge where I need to go next… a way to assess the students just to see if they learned specific content or skills" (I). Chelsey understood that FA involved eliciting student status, assessing what concepts or skills the student has learned, and guiding the decision the teacher took for the next step of instruction.

Chelsey assessed "little chunks" of information usually through questioning at various points of instruction, beginning, middle, or end. According to Chelsey, FA can be based on something “taught recently. You can do a FA five minutes after you present the information, at the end of class, it does not have to occur after giving the information over a long period.” She believed that FA serves as “checkpoints and help guide the instruction.” Chelsey used FA to sense the gap in students learning and uncover misconceptions and used the information to adjust instruction and guide the students to revise their learning. Chelsey explained that this process is vital because

> if they don’t understand something before, we get to the end, I don’t need to get through the unit, and at the end, you did not learn anything. If I can break it down, let’s say you don’t understand this concept, I need you to understand this piece. What do we need to do for you to understand this? What other resources or assistance can we get you to understand this concept? Let’s do this, make sure we reinforce those concepts that you did not learn at the beginning (I)

Chelsey described FA in this quote as an intervention to ensure that she monitors and moves students’ learning towards the goal of the lesson and adjust her instruction accordingly.

Formative Assessment Practice - Eliciting Evidence and Adjusting Teaching and Learning

Chelsey always started her instruction of new concepts by sharing and clarifying the goal of the lesson and informing students of the success criteria. She posted the standard and learning target on one side of the board daily and explained the expectations, to provide students with a destination for which to plan. Chelsey elicited students' present status or where the students were with respect to the goal of the lesson, using a variety of methods, including pre-assessments, pictures, videos, and phenomena. Chelsey asked questions to assess students' prior knowledge and reasoning of the core idea. During a practice activity on genetics, Chelsey asked questions in a Socratic (thought-stimulating) fashion, to solicit multiple responses from students and to assess students evolving understanding with this vignette;

T You have two parents one is homozygous dominant, and the other is homozygous recessive, what will go in this box? (pointing to a square in the Punnett square on the board)

S1 Big P little p

S2 Pp, Pp, Pp, Pp

T Pp?

S1 They all are the same

T They are all the same, so the allele combinations are the same, what can you say about the offspring?

S1 Oh! everybody is going to look the same

S2 They will all look the same

T They are best what the trait represents; that is what you said (responding to S1)?

Chelsey did not only asked questions in a Socratic fashion, but she also encouraged chorus answers from her students to "welcome all viewpoints." She helped her students to reflect on and add to responses from their peers. This was evident during an argument session where groups listened to each other groups present their evidence, asked each other questions and provided feedback, then revised their arguments. The discussions that followed were explicit and reflective in nature as the students shared their understanding of the core ideas and how they could conduct a better investigation.

Chelsey was thoughtful with the assessment task used to monitor students' progress towards the goal of the lesson. She explained that, "any task that I give, there is something specific I want them to learn from it. It may not be everything according to the standard. It may be that I want them to understand something before we truly address the standard” (I)

During classroom discussions, Chelsey tossed students' questions back to them and provided a context for them to build on. She told her students, "I am not the only source of information, I don't know everything" (O). Chelsey continuously checked for understanding by asking questions in a Socratic manner, where students could examine their ideas or evidence they provided for claims and be able to determine whether the evidence was justifiable.
Chelsey created questions that reflected specific skills of the lesson's standard to assess students "understanding of the concepts." However, for questions that were generic as "write three things you learn today, one thing you still have question about," she borrowed them from internet sources.

Chelsey encouraged her students to be self-directed learners and provided opportunities for them to serve as resources for each other. She instructed her students to look up information to answer not only the teacher's questions but their questions and questions from their peers. Chelsey challenged her students on a scavenger hunt activity to search for evidence to support their claims and not limit their search of evidence to only the information provided in the activity sheet. Chelsey reminded her students to "turn in your answers and evidence on your source card" (contains information gathered through research including references). She observed and listened to groups as they presented information, asked each other questions, critiqued each other's work, and provided constructive feedback. She cautioned them to pay attention to "what other people say? What were the comments that they left? What are some of the things that stood out to you? Where can you improve? Because you will need them for your argumentation session". Chelsey gave her students an option on how to communicate their understanding of concepts. For students who "have difficulty expressing their learning on paper," she gave them the option to "present orally," and they stood up and explained it coherently" (I).

After an activity or a test, Chelsey instruct her students to do a reflection in their notebooks. She stated

> I like to get them to talk and write so I can see, so I want them to feel like they can tell me and that I am listening. They talked about how well they do some of the things they did or didn't do. I didn't make the notes on the graph paper as you told me to. Or I did not study. So, I need to come up with different strategies to encourage them to explore, and they are aware of it. So, in the end, it also helps me with my instructional approach. It informs me about what I need to do; it also tells me what the students know and how they learn best. (I)

The purpose of the reflections was to assess students' learning and the context of their learning to provide Chelsey with information to take the necessary next step of instruction.

During the time for individual work, Chelsey called on each student to bring their notebook, and she read them, asked questions, and used evidence of their progress to provide them with feedback to improve on future assignments and advance their understanding of concepts. In addition to the reflections, the next step for instruction implemented by Chelsey included directing students to complete reinforcement activity on Google docs, referring them to the rubric to do corrections, and requiring them to perform error analysis. She also revisited the information most students struggled with and re-assessed the concept. Chelsey used Google classroom to provide students with resources for remediation and activities to reinforce difficult concepts and offered an opportunity to those who could not finish their assignments in class.

This case indicates how Chelsey used feedback to adjust her instruction by providing extra help and re-teaching.

Formative assessment practice: resources

Resources were necessary for Chelsey's practice of FA. She indicated that she employed a variety of resources, including "practice booklets, videos, online practice tests like Kahoot and Poll everywhere, and Khan Academy," to assess students' prior and developing understanding. The choice of resources Chelsey used depended on how much time was available, "sometimes it is just paper and pencil ticket out the door, I am not opposed to old fashion." She indicated that it depends on how the lesson goes, "after I present the information or after the students present the information, and we don't have that much time." She figured out a way to assess her students "quickly." For example, with some resources like students' phones and laptop computers, she always incorporated extra time to be proactive.

Formative assessment practice: challenges

The challenges of practicing FA, according to Chelsey, is student participation, using technology, and time. Chelsey explained that because FA, like ticket-out, the door is an "end of the class assessment, it can be very challenging because students are ready to leave." She is certain any teacher will agree that the "last five minutes of class, the class is over," and it is hard to get something valuable from the students. However, Chelsey tried to refocus her students after the warning bell during a ticket out the door task and said to them "oh no, no, lets' roll it back, let's take this assessment, class is not over." Although she protested, some students picked up their books ready to leave, and "they just put anything down for credit" (I, O). Chelsey realized that she could not put much weight on this information gathered because it "does not reflect students' actual ability or knowledge, and you can't use it to prepare for a future lesson." When students are talking or writing, they reveal important information about what they are thinking and what they are learning. Chelsey struggled with "lack of student participation, which could allow you to uncover what they have learned," and she could use it to make an instructional decision. For example, during a pre-exploratory activity to determine what "types of predictions could be made using Punnett squares," Chelsey calls on a student to share their idea, and the conversation went as follows

T What do you say? Can you tell us what you think will happen?

S1 No

T Just read what is on your paper (no response from student S1)

S2 He got nothing on his paper

S3 He is not ok today

Chelsey called on another student who volunteered to share their idea.

In another instance where students were using chrome books to gather evidence to support their claim for the pattern of inheritance of blood types, two students never opened their computers. At the end of the class, Chelsey sadly explained that one of the "students sat there on her phone all period and did nothing." She provided her students with many opportunities to learn, and some of them refused to.

Chelsey indicated that access to technology was a big problem, it was not always reliable, and she as well as some of her students did not know everything about technology. When students are working with computers, "if something messes up or malfunctions, you have to try to fix it, because if you don't, students will say, I didn't get mine done on time before the bell rung," and it might lead to other issues. Chelsey required her students to return and plug in laptops in a laptop cart (charging station) before leaving, but not all do. On one occasion, the last students returning the laptop called out to Chelsey and said, “they did not charge them,” just before the dismissal bell. Lack of proper maintenance of the laptops explained why students complained and said, “it’s dead,” “I can’t get on,” and Chelsey instructed them to “get a different one.”

Time, according to Chelsey was also a drawback for practicing FA; sometimes when “students get really engaged in certain topics, and they start asking questions, and I have to pull additional resources that I did not intend to use, or they might say if you go here or there, and I start pulling these resources from their suggestions.” Doing all these unexpected but necessary actions resulted in less or no time for FA. For example, when Chelsey showed her student a photo of identical twin girls with different skin colors, one white and the other black, it prompted a heated classroom discussion between students. Some argued that the girls were not identical twins, others argued they were identical, but the evidence provided was not sound. Chelsey intended with this phenomenon to show how the structure of DNA and RNA could lead to the expression of specific traits differently, due to independent assortment of genes during meiosis. One student asked that can “two Black people can have a white baby?” It started another discussion, and Chelsey went on the web and projected another picture, this time of two black parents with a white baby. Chelsey explained that all these issues need to be taken into consideration to make sure that there is “enough time to take the FA because sometimes I do run out of time actually to assess, look at it and then go back and reteach” (I). Chelsey’s lesson plans for the day indicated an assessment, ticket-out the door at the end, but the phenomenon took more than half the class period, and the bell rang before she could give the formal FA.

Three-Dimensional Teaching: Teacher's Conception of Integrated Science Instruction

When Chelsey heard the word three-dimension (3D) during the first interview, she let out a huge sigh to express her feelings and exclaimed, "Oh Lord… It is complicated." She believed that it took time and cautioned that "some people think you can do it all in one day, you can't. When it engages the students, it goes on for a while. So, what you are planning doesn't always work like that." This quote indicated how uneasy Chelsey was with the concept of 3D and her ability to teach it. She continued with the declaration that the process is complex and lengthy. Chelsey believed that it "relates to teaching all three dimensions together, not separately as before." She provided students with "visuals or true stories for them to relate with, do the things that scientists do, rather than me telling them." She described 3D teaching as a method where "students are engaged from the beginning to the end with all the three dimensions." Chelsey explained that when she thinks of 3D teaching;

> I look at the concepts that the students have to learn, the instructional strategy of the teacher, and how we are assessing them. Words that come to mind are things like phenomena. Trying to engage the student into the activity, looking at how this vertically aligns. We are looking at those CCC, how does this align with what they have done in previous science classes… how might literacy and math standards relate to our performance or core science standard? With 3D teaching, you look at all those things and how we can best fit them into this model so students can best understand. I am trying to capture their attention and get them to answer something that is probably going on in real life or something that might arise in the future. I tell them, there are some old and some recent events with questions that still need answers, how are we going to address those

> things that might occur? So, you need to be critical thinkers to get to that point. That’s why we are looking at those different concepts.

Chelsey, in this description, demonstrated her understanding of 3D teaching. It entails involving students in using CCC to connect ideas and apply SEP to DCI and using scenarios that are relevant to students’ interest and life experiences to provide students with the rationale for participating in three dimensions of learning.

Three-Dimensional Practice: Teacher’s Enactment of Integrated Science Instruction

Chelsey stated that she tried to "start with a phenomenon" to engage her students in asking questions and involve them in "discussions, elaborations, and making connections about the concepts to other contexts." The phenomenon provided a context for students to "elaborate on what is going on and how it may affect them," especially interesting scenarios. For example, during the beginning of a new lesson on genetic variation, Chelsey projected a phenomenon about a family of mixed-race couple with White and Black children. Students generated lots of questions and formulated claims from it. Chelsey did not answer students' question but asked them to share their thoughts and to "provide evidence for your claim. What kind of model can you use to justify what you just said?" The following week she displayed another attention-focused opener, a clip of The Maury Povich Show on determining paternity. Chelsey used the phenomena as the anchor in her 3D teaching to elicit students' prior knowledge and to provide context to start the discourses in her classroom. With this opener, Chelsey involved her students in SEP of asking questions and defining problems, developing and using models, and constructing arguments from evidence to elicit their present status and get them ready for the activity that will help move them closer to the learning goal.

Chelsey struggled at the beginning to "get the students interested." She started collaborating with her colleagues because "we wanted to get the students more engaged because we noticed that was the missing piece." Together the team came up with "true stories that were local… it started discussions in the classrooms because students could relate to it." After discussing the paternity video, Chelsey assigned a case study, a "real, local story" about a Georgia woman who killed her two degenerative-diseased sons. Chelsey was referring to the New York Times article; Mother Gets 5 Years for Killing two ill sons, a case study about Huntington's disease. The discussions explored more concepts than what was the target of the lesson, from the possibilities of genetic inheritance to the ethics of mercy killing, and "brought in a lot of questions" from the students. The students made claims and "struggled to construct solutions to their questions." To investigate their claims, Chelsey involved her students in activities that allowed them to obtain, evaluate, and communicate information on their understanding of the targeted standard.

One such activity was the Argument-Driven Inquiry lab on the inheritance of blood type to determine whether all of Mr. Johnson's children were his biological offspring. Students were involved with the practices of making claims, brainstorming experimental designs to test the paternity of the children, and gathering evidence during their investigations to justify their claims and present their findings. Chelsey guided students as they conducted their research, collected data, displayed their information on a white-board, and shared the evidence that supported their claim. She provided a rubric that accompanied the tasks as criteria for success and for students to give anonymous and honest feedback to their peers, to encourage valuable evaluation of each other's work. This example also illustrated how Chelsey involved her students with CCC of cause and effect and patterns in helping them link the practices to the core idea of genetic inheritance. Chelsey involved her students in activities like case studies and argumentation-driven inquiry because these activities addressed each of the dimensions at some point. After each task, she revisited the targeted standard and reviewed the vital point with the students. She directed her students to visit google classroom for more practice and gather information. Most of the class period was spent discovering a phenomenon, asking questions, making and supporting claims, and little time spent on lecturing. Chelsey's practice of 3D is summarized in Table 4. as shown below.

Table 4.
Three-dimensional Science Practice

Lesson Title: Mother Kills Three Sons out of Love

Standard: Obtain, evaluate, and communicate information to analyze how biological traits are passed on to successive generations.

	Action	**DCI**	**SEP**	**CCC**
Pre-Instruction	Plan from standard	Inheritance	All related SEP	Patterns Cause & effects
Beginning	Reviewing concepts Phenomena (genetic disorder-Huntington) Class discussions	Inheritance	Asking questions (making claims and proposing solutions)	Cause and effects
Middle	Identifying and researching on a genetic disorder (individual/group) Group discussions	Inheritance	Planning and carrying out investigations	Cause and effect
Last	Continuing research and preparing presentation Homework	Inheritance	Obtain, evaluate, and communicate information	Cause and effect

Chelsey involved her students in cross-disciplinary concepts, including concepts in previous and subsequent grades, to enhance students' understanding. During classroom discussions, Chelsey and her students brought in ideas from "other disciplines and different aspects," to make sense of the topic. She wanted her students to reflect, for instance, "how does what you are leaning in another class relate to what you are learning here?" to help students make connections with the concepts they were learning. Chelsey saw promise in the future of 3D teaching as she declared, “I think the 3D learning model is beneficial.”

Chelsey talked positively about the role of collaborative planning in helping her navigate through this new way of teaching. She expressed that when planning together, looking at the different dimensions, “each one of us brings in several different ideas… one might suggest a video; another might say it should be under five minutes. Another might suggest an article that is simple but informative and reinforces the concepts.”

Three-dimensional practice: challenges

The challenges Chelsey faced with teaching 3D is the push back from students, resistance from teachers, and lack of time to meet and collaborate. Chelsey explained that “everyone is not doing it and is confusing to the students.” Chelsey tried to follow the requirements of the 3D model, and she believed that when students see it in other classrooms, “they can relate” with it. However, the lack of full implementation from all teachers contributed to Chelsey’s “students complaining, and they say, why don’t you get up and teach, can you just lecture today?” Chelsey shared her frustration with students telling her, “you do not teach” and comparing her to those teachers who lecture. Chelsey believed that if all teachers were on board, it would make a difference.

> If we correctly utilize this, everybody in their instruction, then It will be more comfortable for the students because they know the structure, the order how we are going to function in an educational setting. But if everybody is not using it, then the students don’t know it.

Chelsey explained that because not all teachers are participating in 3D teaching, “when you are trying to implement it, the students feel like it is a trick, that you are trying to fail them.” Given also that her students are not successful at the beginning, they "are ready to give up, they say I am failing because you are teaching this kind of way." Chelsey tried to convince her students to buy into this new way of teaching by explaining to them that

> talking to you is not teaching… for me to talk to you the whole class period and not with you, not having a discussion, I don't feel like you are learning. Studies show that you are not learning just because I am talking to you. (O)

The irony Chelsey lamented is that the "same students who complained about you not teaching if you try to talk for more than five minutes, you lose them" (I). Chelsey revealed that she must negotiate with her students if she sees the need to explain a concept. "I tell them to give me three minutes, and I will leave you alone. I am trying to get them away from the idea that you should stand in front of the class and teach" (I). She encouraged her students to "talk with your partner, your group members," and not to depend solely on her. She explained that "I am not the only source of information, you have more information between all of you, and you just need to tap into it" (O).

Chelsey tried to foster student learning by giving them tasks to explore and to figure out for themselves, but her students fight back with, "why do we have to do this, you have not taught us anything yet." Chelsey explained to the class that

> I am not going to tell you anything yet because there are problems that have not been solved yet, there are problems that will come your way and if I tell you the answer now, how are you going to learn how to solve the problem? You don't know how to solve a problem; you are just listening for an answer. When a question arises, I need you to think through because the answer will not be there, or someone may not have the answer to give to you. (O)

Chelsey encouraged and solicited students to be owners of their own learning and to think like a scientist.

FA3D: Gaging Students' Integrated Science Knowledge and Adjusting Instruction

Chelsey assessed students integrated science learning using a variety of FA approaches, including questioning students' reasoning, observing their performances and the skills they used, and evaluating their writings. She explained that she had been "doing constant checks just to see" what her students know. She provided assessment tasks that required students to do "a lot of reflections, as a type of FA of what they had been learning." For example, Chelsey involved students with scenarios that provided context for her to assess how students employed SEP (asking questions, making claims, carrying out research, and constructing solutions), CCC (observing patterns and cause and effects) to connect the SEP and the DCI. In one such task, Chelsey presented her students with a phenomenon, a short U tube video on paternity to assess how students made claims, provided evidence to support their claims, and suggested ways to investigate paternity. Chelsey, in her assessment of 3D learning, used SEP tools to elicit evidence of students' knowledge and uncover misconceptions through asking questions and communicating information. During the use of phenomena (involving skin color and paternity), she observed and asked questions to assess how students made observations, asked questions, construct claims, and answered their questions or suggested investigations to answer their questions.

Chelsey encouraged classroom discourse and peer evaluation. She instructed her students to "talk with your group members, talk with your partner. What do your peers say about what you are doing?" (O). Having that type of discourse in the classroom environment where "yes, we can talk about it and if you don't know it, let's ask somebody, let's have those meaningful discussions in class instead of talking about who is going to twerk best or who is going to whose party" (O). Chelsey activated students as resources for self and each other in the assessment process. She encouraged students to ask questions and build on each other's ideas, and she modeled how students could act as resources for each other. She provided opportunities for more student-talk to "try and cut down on those types of discussions and more into let's have a scientific discussion; let's learn something."

To structure her assessment task, Chelsey used the "UBD (Understanding by Design), backward design" model to guide her. "looking at the outcome and constructing questions to embed in the lesson" (I). Her questions ranged in rigor from "baseline questions that could be level one and gradually build them up to level three questions" to meet the standard and the needs of students with a spectrum of abilities. The purpose of these questions was to "check for understanding or whether students were able to apply that knowledge into another concept." Chelsey had questions in place or anticipated certain questions from students during classroom discussions to address misconceptions. She was purposeful with her questions, "have them at the back of my mind, put them in the lesson plan; that's what I have been doing more" (I).

She shared that working with colleagues facilitated the process. Chelsey and her teammates brainstormed together to come up with "standard-based questions" (questions that were diverse in content and format and contained small aspects of each dimension). In their planning, Chelsey and team members considered ideas about how to formatively assess 3D learning, the DCI to drive the lesson, the SEP students would use to understand phenomena, and the CCC to connect DCI and SEP. Chelsey said that she compiled "the types of questions I am expecting the students to ask and these are the ones I will ask." Chelsey believed that knowing what she is "going to do and anticipating what the students are going to say and the type of discussions we might have is good." Here Chelsey revealed how she identified and used DCIs as a central organizing concept for classroom activities and her assessment tasks.

After every assessment, Chelsey required her students to "do reflections of two to three sentences, easy to read, and it's beneficial" (I). All students automatically earned a grade for their reflections, "but they were not doing it just for the grade because some of them don't do it." Some of their reflections read, "I didn't do this; I need to come to class often." Chelsey could see how the students felt about 3D, and she noticed that her students were honest with their comments in their writings. "I want them to reflect because they are putting the responsibility on themselves, and I like that." The reflections gave students the opportunity to express in their writings how they felt about the standard, whether it was addressed adequately, questions about the concepts, how well they learned the concept, and whether they could master it on a quiz. Chelsey said that I "wanted them to reflect on their personal learning…to feel like they can tell me and that I am listening" (I). The reflections also allowed students to be "more open with their writings, and they were asking more questions." Chelsey said that students were not shy to ask, for example, the meaning of words, "which I am glad they do instead of the [let's keep on moving attitude], so they can get to that level of understanding before moving on."

Chelsey also used the ticket out the door and surveys for students to reveal their understanding of the concept(s) discussed for the day. The students also completed an error analysis to "make corrections on questions they missed, to justify their reasoning with evidence, and to talk about their misconceptions and reasons they missed the questions." Completing the error analysis encouraged Chelsey's students to take ownership of their learning, and she can also assess how the students used SEP and CCC in their explanations.

Chelsey revisited missed questions in her warmups and future instructions. She explained that “I am still going to address that concept in subsequent lessons, and I know I will keep emphasizing the concept because they did not get it” (I). For example, Chelsey requested students to “take pictures of your ADI white-boards an upload into Google classroom,” and she provided comments on them. The following day she shared with students the strength and weaknesses of the information on their boards and what to watch for when completing their reflections that followed the presentations.

Formative assessment of 3D: challenges

Chelsey indicated that her primary challenge with assessing 3D is “doing the backward” design. Trying from the beginning to design the task or “come up with questions that address all three dimensions and the pre- and post-assessments” that meet the requirement for the goal of the lesson. Starting with the end in mind is important because “we do a pre-assessment and use that to focus on the lesson.” Chelsey used the pre-assessments to elicit students’ status and misconceptions they might have. Her assessments were multi-components as she made sure one of the questions address each of the dimensions to include all three.

Chelsey’s Summary

Chelsey described FA as anything that helped her checked for students understanding to see what they knew and had learned, and that gave her information to make an instructional decision. Formative assessment accesses information over a short period of time. Chelsey used assessment tasks created based on the goal of the standard to elicit students’ status and gauge their learning. She asked questions that were thought stimulating (Socratic pattern) and used questions that addressed the elements of the standard. She encouraged her students to be resources for each other and to practice self and peer evaluation. Chelsey experienced some obstacles implementing FA, including lack of student participation, time, and problems with technology.

Chelsey's conception of 3D teaching was that it was very complicated. She believed that integrated science instruction took time to complete, and it was naïve to think that it could be done all in a day. Chelsey is certain the process involved teaching the three dimensions together, starting with a phenomenon to keep the students engaged from start to finish. Her students participated in the SEP of asking questions during discussions, and in using CCC to make connections and link SEP to DCI. She used the phenomenon to elicit students' prior knowledge and to provide a context to start the discourse in her classroom. Collaboration with colleagues was helpful for 3D teaching, but Chelsey complained that this did not happen regularly. Chelsey used case studies and argumentative inquiry activities to involve her students in the elements of the three dimensions. She provided rubrics with these activities as criteria for success or to achieve the goal of the lesson. These activities were selected based on the requirement of the standard. Chelsey struggled from lack of full participation from students, they were apprehensive of this new way of teaching and feared that it was a trick to flunk them, given that not all teachers were practicing 3D teaching yet.

Chelsey assessed students' understanding of 3D learning using a variety of means. She elicited students' prior knowledge or status using warmups, hands up, and pre-assessments. She constantly checked how students were using and applying the elements of the three dimensions during their participation in the case study and argumentative activities, through questioning their reasoning, observing their performances and the skills they used, and evaluating their writings. Chelsey used assessment tasks that were standard based to monitor students' progress along a continuum of developing performance expectations. She encouraged self and peer assessment of student work and required her students to provide constructive feedback to each other. Chelsey involved her students in completing reflections and error analysis to be informed of their strengths and weaknesses and take ownership of their learning. Her challenge practicing FA of 3D was coming up with questions from the beginning that reflected the standard or doing the "backward design."

Case 2. Paul

Paul is the most senior of the three participants and a male. He has been teaching science going into his 17th year. Paul has a bachelor's degree with a major in French and a minor in science. He earned his certification through an alternate route. He started his teaching carrier as a French teacher but switched to science after taking some graduate science courses and passing the state board certification exam. Paul has been teaching 9th and 10th grade biology at his present school since 2003.

Teacher's understanding of Formative Assessment- Constant Checks

When asked about his understanding of FA, Paul explained that "for me in the past, it has always been a test." Paul described FA as "anything where I can get an idea of how well my students have learned the topic… constantly checking where they are." With this shift in thinking, FA to Paul "may look a variety of ways. It may be a straight-up multiple-choice test, a writing question here and there, arguments, discussions… completion of a contractual amount of work, or a bell-ringer, ticket out the door." The goal of these strategies was to "elicit students' status, evidence of their learning, and uncover misconceptions." Paul explained that once he completed a concept, "anything I give them, and I get that feedback from them, I can use the information to decide if I need to remediate or not, that is formative." According to Paul's description, FA was anything that monitored or tracked students' progress towards the goal of the lesson, something that elicited student status and gave him information to take the next step on instruction. Paul understood that FA helped him diagnose what students came in with, what students were learning, and gave him the information to make an instructional adjustment and steered students on the right path towards meeting the goal of the lesson. Paul believed that FA is a process that "brings clarity to what we do."

Formative Assessment Practice: Eliciting Evidence and Adjusting Teaching and Learning

Paul identified and clarified the goal of the lesson for the students at the beginning. As he puts it, "if they don't know where to go, and I don't make it clear where they are heading, then that's going to be a problem" (I). It was necessary to identify and share success criteria to inform where the student was going. Paul posted the goal of the lesson or learning target on the board daily, so he reminded students of where they were going daily. He also posted warmup questions below the learning target to tell students of its importance. It assessed where the students were in their learning with respect to the standard, and it gave Paul a sense of where to start his instruction.

Paul asked questions in a systematic, Socratic pattern using probing questions to elicit thoughts from many directions and analyze complex ideas. This technic helped Paul uncover students' assumptions, misconceptions, and their strength and weaknesses. The vignette below portrays Paul's questioning technique to assess students' status (where the student is) and understanding of the concept of inheritance from an episode of the Maury show.

T How do we know for sure he is not the father, because Maury says so?

S No, DNA test, sperm

T So, DNA test, what does it do?

S Test the blood

T Blood test hmm. Would sperm give a complete DNA profile?

S No, yes, because there are a million of them

T If I have a million sperm, do I have a million DNA?

Ss No you only get half

T You only get half, that's what I am talking about, you only get half in sperm

Paul used this vignette to assess whether students understood inheritance DNA from both parents (half from each parent). He challenged their thinking with his questions, compared their ideas with the goal of the lesson, provided feedback to them, and encouraged peer and self-assessment.

Paul responded to students' questions with a reflective toss. He explained that "If they ask me a question, they get a question back from me. There is not an answer. I don't give answers to them. I am going to give them a probing question that will help them come up with the answer themselves" (I). Paul rephrased and summarized what students said and stimulated comments from peers. He interpreted students' comments by stating, "let me put together what you just said" (O). In that way, the student and his or her peers can process the information. Paul, in these situations, was interpreting information by repeating and clarifying students' questions and comments, to show where the students are in their learning. Paul encouraged students to exchange ideas with each other, "bounce your ideas off each other" (O). He also stimulated students to take ownership of their learning and to become resources for each other. When a student asked a question, Paul reflected and said, "how can I get these kids to answer this question themselves" (I). The excerpt below further provides an example of this practice by Paul.

T I want you to analyze your mistakes with your peers

S You mean I am going to look at this in class to see if they think it's right or not?

P Yes, put your faith, trust in your teammates because they are going to be the one to read your paper.

In his practice of FA, Paul modeled for students how to act as resources for each other. He explained to them that "I am not the only one in the room, or my knowledge about the topic can't be the only resource" (O). Paul encouraged students to consider other options because "not all people will learn from the one resource I give." Paul fostered critical thinking and discouraged dependence. When students tried to use their phones to obtain answers rather than think through the problem, Paul discouraged them by stating, "don't bother Googling; I wrote the questions myself." In these examples, he engaged his students as master of their learning. On one such occasion, Paul challenged his students' thinking by requiring them to self-evaluate their work, stating, "I will return them to you to crosscheck before I grade them" (O). Paul provided students with the means to achieve the goal of the lesson or how to get there. During an interactive discussion of different blood groups, Paul gave student feedback as constructive criticism as followed, "I love the way your claim sounds. Your justification statement gives the reason why. You brought in this data; great, I like what you just said" (O).

Paul believed that during instruction, he does not know yet how he is teaching or whether his students are getting it, "until I give that FA, there is no reality to it… I do not know the answer if they know it or not. What it does is that it brings clarity; you can have those real conversations at that point" (I) and take the next step of instruction.

Paul indicated that assessing students is not the end of his FA practice. He must take the next step. After eliciting students' status and providing a road map on how to close the gap in their learning, Paul used the feedback to determine whether students still struggle with the concepts. His next step for instruction may include requiring students to come for a tutorial or analyze their mistakes. With one student he explained that "I can start you with something that seems like elementary school, then maybe that is where I need to start you" and with another student, he explained that "you have already shown me that you know this, so I need you to continue writing that paper (investigative report)" (I). Paul also used feedback (data from quiz) to group his students into mixed abilities during activities.

Paul gave his students the opportunity to advance their thinking by completing "an error analysis activity with each question you missed. We are going to find out why you missed the questions" to clarify misconceptions and foster understanding. This quote indicated how Paul acted on the evidence of student thinking and helped his students close the gap in their learning or moved them towards the learning goal. Paul also gave students opportunities to revise poorly completed assignments and resubmit them based on directive feedback.

Formative assessment practice: resources

Paul assembled resources from different sources to assist with FA practice, such as from the "Georgia department of education website, the curriculum standard, even from your colleagues." Paul and his colleagues planned their lessons together to come up with questions and activities to examine how students develop an understanding of concepts. He asserted that "collaboration is key to sharing those resources and pulling them into one place." It is important because "my knowledge about the topic can't be the only resource… You have to make sure you have and could give students every resource available." Paul emphasized that the standard was the main resource and guided what he did or what students should know and could do from the beginning to the end of the lesson. He explained that because the students "might not always get it with these resources, so I need to take you back to square one… and that is the standard."

Paul summarized his practice of FA with this excerpt

> I am practicing FA when they are writing their rough drafts. If you don't get it, that's why it is a rough draft, and this one I am not reading it, no, your peers are going to read it first. I would not read it until you complete it. I am going to make the suggestions at the end, so; I am going to be purposeful. The assessments are going to be double-blind. I am going to make sure you feel comfortable with this.

This assessment afforded the student and teacher with information on where the student is, so the teacher can provide the appropriate support for students to close the gap in their learning.

Three-Dimensional Teaching: Teacher's Conception of Integrated Science Instruction

Paul described 3D teaching as a way of teaching where you teach all the related concepts from each dimension together. Paul felt that the process was "tricky." He always related what he was teaching to "science overall so students can see the big picture. If they can't see the big picture, then they can't put things together" to understand science. Paul explained that it is necessary to teach all the sciences together, "as a biology teacher, I have to teach the rule of chemistry, I have to teach the rule of physics… It's important to teach all those concepts in the curriculum that supports the standard, that is the way it should be every day" (I) for 3D teaching. Paul believed that for his students to learn science, he must do at most 35% of the work. He told his students that "I am not the one learning; I already know it" (O), and explained that "for them to learn it, they have to do it. Do something about it. Its' not just me talking" (I).

Three-Dimensional teaching is a "way of exposing students to all the dimensions so, at the end, they can say how one concept relates to or influences the other." Paul demonstrated his understanding of the integrated nature of 3D teaching with this quote.

> Doing an investigation leads into an experiment, scientists are trying to solve a problem that's what makes it different from the other areas. In math, the teacher knows the answer

> before they ask the question. In social studies, there are debates about the causes of war and other historical events, but for the most part, all that history is exhausted. You can explain the reason for the civil war, the reason for the financial collapse. In English, if the subject does not match the verb, you are going to get points taken off. So, in all those three areas, the correct answer is pretty much there. We differ, in science when the professional goes into the field, they are trying to solve a problem without a solution yet. Keep that in mind during our work session, about trying to find an answer to a question without one. I can relate concepts in different grades especially in Biology. I must teach the rule of chemistry, of physics, how life forms emerge, if the life forms are successful, by those laws, then how does it relate to sub life, below life. When I start seeing students state reasons, justifying their reasoning, quoting the laws of thermodynamic or laws of energy conservation, then they are getting it.

This quote demonstrated Paul's understanding of 3D teaching as grounded in integrating the DCI, SEP, and CCC, and involving the students in this kind of learning.

Three-Dimensional Practice: Teacher's Enactment of Integrated Science Instruction

The practice of 3D according to Paul was doing "something different from what I had done in the past." Paul described that in the practice of 3D, he does "all sort of things, starting with some phenomena, something they notice, something relatable that grabs their attention, and we have to build on experience." Paul implied in this description that 3D teaching had multi-components and that the phenomenon must be related to the student's interest and life experience. The phenomena also provided a "rationale for why the concept was important" or relevant to student's life. The phenomena thus must be chosen with care because students had to have noticed it before, for it to grabs their attention. It is something that gives students the prior knowledge or “provides the context, something that gets them prepared for the ADI, for their CER” (ADI is an argument-driven inquiry and CER is a claim, evidence, reasoning).

Paul started his lessons with careful planning from the standard as evident from his lesson plans. Paul said that during collaborative planning he gathered related questions that accompanied the phenomenon and the subsequent activity to guide students’ discussions and exploration. To grab students’ attention and begin the conversation in the classroom, Paul provided a clip of the Maury show about paternity as an introduction to the concept of inheritance. He played the clip and stopped just before Maury revealed the results. The students were anxious to know the results and begged Paul to continue the video, but Paul reminded them of the thinking process by asking, “what do we do when we have a question to answer? Is he the father? The students responded in chorus with one-word answers of yes or no. Paul exclaimed and said, “wow, wow, wow. I am sorry, I did not realize it was called C. I thought it was called CER. If all you are doing is making a claim, that is just noise.” Paul calmed down his students and asked them to take a moment and think through the scenario. He encouraged them to “make sure you can back this up. I want to hear your reasoning.” With this introduction, Paul prepared the students for the next activity involving argumentation where students would make claims, develop proposals, design and carry out investigation to solve real life problems. Paul used the ADI activity to help students used relevant CCC to connect the SEP with the core idea of the problem under investigation.

Paul motivated his students and encouraged them to think differently with this new way of learning. He told them during class discussions that the next discovery is on them because “people my age hasn’t figured it out yet. It’s your turn” to lead future inventions. He inspired his students to take risk in their thinking and bring ideas from other domains to make connections with what they see because they are the future and added that you have to “think critically; we will need some heroes, some science heroes.” Providing his students with the rationale to “do science” (3D learning) according to Paul, will assure that they “never ask the one question, when am I going to use this ever?” because the reasons will be obvious.

Paul strived to provide equal weight to the three dimensions in his practice of 3D instruction. However, he is convinced that content is the core of 3D teaching. “I got to go to content. Because the content itself is what is going to drive the questions and everything that is on the test.” Though he recognized that for the lesson to be 3D, he must “also teach how to think like a scientist… the third is certain themes that run through all sciences, CCC.” Paul said that he made sure that the students “hear something from me daily in all those three areas and not being so content heavy” (I). I observed this pattern in the classroom. Paul involved the students in SEP when they were asking questions, making claims, and carrying out investigations. He involved students in CCC during a warm-up activity following the phenomenon where students used cause and effects and patterns to justify their evidence. For example, to teach students how to use patterns to connect the practices with the core idea, Paul used an activity where students explained genetic inheritance. First, they explored how meteorologist predict when a tornado or hurricane will hit and the path it will take based on patterns, and how cardiologist use patterns of heart bit to determine a healthy heart. Then the students moved on to the concept of patterns of inheritance based on observations they made. Paul recognized that “it is important to teach all those concepts,” to allow students to “analyze their performance to build on their knowledge and minimize misconceptions.” Paul believed it is necessary to provide opportunities for students to learn science, that is to apply what they are learning beyond science classes, and that is “3D learning.” Table 5. below provides a window into Paul’s practice of 3D.

Table 5.
Paul's Practice of 3D

Lesson title: The Royal disease from Tainted Blood

Standard: Obtain, evaluate, and communicate information to analyze how biological traits are passed on to successive generations.

	Action	DCI	SEP	CCC
Pre-Instruction	Plan from standard	Inheritance	All related SEP	Pattern Cause & effects
Beginning	Warmups Phenomena (genetic disorder- Hemophilia) Class discussions	Inheritance	Asking questions Construct and use models,	Pattern Cause & effect
Middle	Constructing pedigrees and determining outcomes Group discussions	Inheritance	Analyzing and interpreting data Constructing explanations	Pattern Cause and effect
Last	Gathering evidence based on Mendel's law of dominance to support results Homework- identify disorder & use pedigree to explain inheritance in google classroom	Inheritance	Obtain, evaluate, and communicate information	Pattern Cause & effect

In his practice of 3D, Paul was confident that the curriculum standard was necessary to guide instruction; it provided a road map to follow. Paul's approach to implementing 3D depended on what the assignments were. For example, Paul said when using the ADI process,

> I am going to be teaching in several different ways. I am going to give them the intro; I am going to avoid front-loading as much as possible, give them the bare bone information they need, to help them start with their proposals. I sit back and watch them do the processes of science and see how well they are picking up on that.

Paul described here how he guided students with a lesson using a phenomenon along with other activities that allowed the students to integrate the three dimensions and to take control of their learning.

Three-dimensional practice: resources

Resources were also necessary to help Paul with the implementation of 3D teaching. The curriculum standard was the main resource for Paul as he used it to guide the lesson. Paul used google classroom to make sure the students "know where we are in the process." Students can see that "I am going into what scientist do now, I am not talking about biology now, and they can guess where I am going with all that." The county also provided some resources and Paul has a little Vernier poster on his wall so, "I know what the SEP is, and I point to it when I talk about it." For content resources, "it could be everything from the textbook to online sources."

Practice of 3D: challenges

Paul's challenge teaching 3D is being able to "remember to do it. To not feel like I am here for the content" only. Paul acknowledges that "because of the pressures we get to make sure we cover all these contents; it is so easy to be bogged down on content that you forget" the other dimensions. Between the two non-content dimensions, "the one I have to make a concerted effort to remember is the concepts that thread through all sciences regardless of what you are doing, the CCC is most difficult to remember." When planning his lessons, Paul integrated the dimensions. Another major challenge is time. When teaching 3D,

> You must be extremely flexible with what you are doing in your classroom. You might have it all planned, and then a student asks you a question, or a student takes you in a direction you haven't thought of. You must be quick on your toes to see where I can incorporate this. How can I answer this question and approach it from a CCC, cross all science? How can I approach this from how scientist think? And how can I answer that question using biology to bridge that gap without making it content heavy? In doing the unplanned, time is of the essence.

Time is necessary for the practice of 3D. To compensate for lack of time, Paul assigned unfinished class work for homework and provided additional support on google classroom to help students meet the goal of 3D teaching. The following day Paul started with the homework and asked, “any questions about the homework? You were taking what you got from the video to make connections and build that pedigree.” Paul hoped students completed their assignments to allow more time for classroom discussion. Given that the warm-up activities that assess how students were applying their understanding of concepts consumed almost half the class period. For example, when Paul guided students in developing their pedigree to explain their thought and make connection to genetic inheritance, it took almost the entire class time.

FA3D: Gaging Students’ Integrated Science Knowledge and Adjusting Instruction

Paul assessed students’ understanding of 3D learning using assessment tasks that were constructed using the standard as a guide. The tasks assess students’ performance and skills, how they used SEP to make sense of phenomena and how they used CCC to make connections between the SEP and the DCI. The evidence from these assessment tasks was used by Paul to decide the next step of instruction. According to Paul, the “standard drives what we do.” For example, Paul was teaching from the standard which demanded the students to “obtain, evaluate, and communicate information to analyze how biological traits are passed on to successive generations.” To assess an element of this standard, Paul involved students in activities using mathematical models to predict and explain patterns of inheritance. One such activity was on color blindness which involved students looking at various pictures with different color patterns to identify what color they were blind to and to explain how the traits were inherited differently. The students then constructed Punnett squares (model) and used the patterns observed to determine the probability of inheritance. This generated classroom discussions and lots of questions. An example of a short exchange in the classroom went as followed

S if people are color blind how does that affect their driving?"

T do you notice anything about the streetlights? Paul tossed

S They change colors?

T Oh! How?

S1 green, red…

S2 patterns

T Patterns, that's what it is

Paul walked students towards the practice of self-discovery. He explained that because the lights were sequenced, drivers who were colorblind may use the changing pattern in traffic lights as cues. Through this activity, Paul assessed what students understood about inheritance and how they used it to predict results.

Paul specified that the activity might involve students in "developing models, justifying their evidence or providing claim, evidence, reasoning to concepts, which varies based on what they can do, and I have to differentiate for some." Paul's assessment tasks assessed each of the dimensions and different aspects of more than one performance expectations. He stated that "not all my questions are about biology. Some questions must be about the process of planning an experiment. Or sometimes if I ask a question about structure, then the answer choice would have a functional value to it" (I). This assertion was observed in Paul's classroom instruction when he asked a question to the class during a warm-up activity. Paul said, "why does a harmer have two heads? why is a harmer shaped differently from a wrench?" The students responded, "because it has different purposes, different functions" (O). He used this authentic example to show the relationship among different blood groups and why they have different antigens on them (because of their different roles). Paul indicated that he disguised questions sometimes and asked them as Biology questions, “but I know that when I am writing that assessment, it is not straight content, because that is not how the national exams or test is going to look like” (I). Some sample questions that Paul asked his students included

- What do scientists do when they encounter a problem?
- Study the four pictures on the board. Which could be blood group A, B, AB, and O? you must justify your answers because that is what we do in science
- What are the major blood types and how are red blood cells from these blood types differ from each other?
- What is the offspring’s phenotypic ratio when two heterozygous plants for height and pod color cross-pollinate? What model will you use and what law, to justify how the alleles are inherited? (O)

Paul formatively assessed students’ 3D learning using several means. He primarily assessed through asking questions to students or from the question’s students asked, in whole class discussion. He observed and listened to students during their investigations and argumentation activities. Paul used concepts that students struggled with as warmups for the following day, to assess their learning and build on with new ideas. He also gave quizzes at the end of covering a concept.

Paul specified that after every assessment, “I do data analysis, look how well they perform and see where I need to remediate” (I), to take the next step of instruction. For example, Paul used information from quiz data to group students into activity zones (station) during an activity session building pedigrees and Punnett squares to determine possible offspring. He also gave students the opportunity to resubmit assignments that he returned for revision to address the standard and its vocabulary (because students used “big B” and “little b,” “little l” and “little l” instead of heterozygous and homozygous respectively and did not interpret their models). After the warmup activity, he told the students, “I will return them to you to crosscheck before I grade them” (O) and returned the previously completed assignments to the students.

Paul thought that he was not confident at this time on how to formatively assess 3D learning. He confessed during the second interview that, “I am trying. Hopefully, I vary my assessment such that they have all different aspects of what might be in the standard (is 3D). You have to know it across the subject” (I). This description indicated that Paul understood that to formative assess 3D learning, the assessment must be multi-components, variable, and not focus solely on content.

Formative assessment of 3D teaching: challenges

When asked what the challenges were for formatively assessing 3D learning, the answer he gave was, “lack of student participation.” Students felt like “there is no repercussion for not doing work; it’s another test so what? Students are not motivated to want to perform to the best of their abilities.” His concern was validated during an interactive discussion on genetic inheritance that came after the phenomenon on Maury. Paul said to the students, “for those of you that bother to answer out loud thank you. Those of you who don’t; I will check it in writing when I see those assignments.” At the end of class, Paul shared that, the same students who participated in class are the same ones who cared about their work. He specified that the behavior is more evident during a quiz, “they crisscross the answers, or you look at their online assessment, and you see three minutes for 30 questions” (I). Some students refused to answer the constructive response questions and others avoided the explanation part making it difficult for

Paul to assess how they used SEP and CCC respectively to make connections and figure out the natural world. Paul is saddened about students' unwillingness to self-invest in their education but is powerless. "I can't do anything but fire you up; you got to do it yourself" (O). However, Paul tries to motivate his students but,

> they try to do whatever they can to fight you to learn. Those are the wants that need that motivation. If it means to reach out to mom, Eh, for me to have a partner at home. If I have to think about something entirely different. It has been days that I have gone home totally confused, and I need to take some time to think about how I am going to get to this one (I)

This quote shed light on how Paul struggled to elicit students' ideas, to involve them in using skills to perform standards-based tasks, and assess students integrated science knowledge.

Paul's Summary

Paul's explanation of his thought of FA suggested a shift in thinking from using "straight test to using anything" that gave him information on how students were doing and to make an instructional decision. He used test, arguments, bell ringers, TOTD, and completion of the task to elicit students' status, evidence of their learning, and any misconceptions. Paul identified and shared the goal of the lesson or success criteria at the beginning. He always posted standard and learning target on the board as a reminder for students. Paul asked a lot of questions, tossed students questions back to them, and solicited chorus responses to elicit thinking, reflections, and multiple ideas from students. He also involved students in self and peer evaluation, modeled how students could become resources for each other, and challenged them to become masters of their learning. Paul used the feedback information to adjust student learning and his instruction. For this next step, he directed his students to complete error analysis and attend tutoring, retaught the lesson, revised tasks, and set up student groups. Paul assembled resources from different web sources and from colleagues to help with the practice of FA.

Paul's understanding of 3D teaching was that it is a way of teaching by integrating all the related concepts from each dimension. He believed that it was necessary to put everything together so students could see the big picture and foster their learning. Also, for his students to learn science, they had to be the ones doing the talking and thinking through asking questions. Paul practiced 3D teaching starting with a phenomenon which must be chosen wisely to meet the needs of all students and to initiate and fuel the conversations in the classroom. Paul created and used activities with elements from all three dimensions as specified in the standard, to include students in integrated science learning. Paul involved his students in activities such as the ADI where they perform claim, evidence, and reasoning to prepare them for scientific explanation writing. The curriculum was the leading resource that guided Paul's practice of 3D teaching, but he also used wall posters of the different dimensions as reminders. Paul's challenge in practicing 3D teaching was to remember to incorporate all three during instruction.

To monitor what students knew and their learning of the three dimensions to adjust instruction, Paul turned to the standard for guidance. He structured questions and assessment tasks that would assess what students knew and could do from the start. The assessment contained practices of all three dimensions and aspects of more than one performance expectations. He identified and shared the learning goals as well as the dimensions with his students at the beginning. To elicit students' status, he used relatable phenomena and questions to gauge what they know and the misconceptions they had. The activities that students participated in assessed how they developed and used models, asked questions and constructed solutions, made claims and provided evidence and reasoning to concepts, and used patterns to make connections. He also employed non-standard based activities to support and build students' skills towards the main activities. Paul used feedback information from analyzing assessment results of students' 3D learning to remediate their learning and adjust his instruction accordingly. The challenge of formative assessing 3D learning according to Paul was students' unwillingness to share what they knew.

Case 3. Andria

Andria is one of the female teachers and with the least teaching experience. She is in her fifth year of teaching, holds a master's degree in secondary science education, and attended a traditional teacher certification preparation program. She is currently teaching 9th grade biology and has for the last four years.

Teacher's Understanding of Formative Assessment- Snapshots

Andria described FA as a "quick quiz, ten minutes weekly FA, students take it to make sure they understand the lesson's standard and elements, to gauge what the kids are learning." Andria believed that it was necessary to identify and communicate where the students were going. She demonstrated this in one of her quotes where she said, "we start with the standard, and it guides what we do. I inform them at the beginning what the assessment will be, as I explain the standard. I tell them this is what you will need to be able to do at the end." Andria also recognized the need to elicit student's status or where the students were, using variable assessment approaches like "ticket out the door, thumbs up or down, sticky notes, and asking questions." Andria believed that anyone of or a combination of these approaches "gives you a quick snapshot of your children, where they are immediately, instead of waiting until the end" (summative). Information on student learning gathered from the formative assessment activities "immediately gives you evidence to either advance or go back and reteach." Andria explained during the first interview that practicing FA can be "as simple as asking students to raise their hands. Giving students warm up at the beginning of class with at most five questions to complete individually."

Andria summed up her understanding of FA when she stated that it "is a snapshot; it serves as a quick reinforcement of students' learning and my instruction." This description of FA by Andria demonstrated her understanding of the role she plays in the formative process. She identified, clarified, and shared learning goals with students, and used the results to track the progress of learning and teaching

FA Practice: Eliciting Evidence and Adjusting Teaching and Learning

At the beginning of each class period, Andria had three to four warmup questions on the overhead projector ready for students to begin working on as they entered the classroom. To elicit what each student knew, she encouraged them to answer the warmup individually. The whole class then reviewed the questions and "students explain why they missed it, why they got it wrong and why it's wrong" (I). During these icebreaking activities, Andria elicited student's status and provided them with feedback on their learning. She walked to each table of four students and assigned a percentage to each student's completed warmup without telling them which question(s), they got wrong. (she collected the warmups sometimes). Before she reviewed the questions, she directed students' attention towards the monitor, "if you did not make a hundred, I need you to make sure you correct your answers…now you will know which one you got wrong" (O), take note of them. To get an accurate understanding of what students knew, Andria did not reveal the right answer immediately so each student will have an opportunity to respond. Andria then reviewed the answers with the students. She called on a volunteer to read the question out loud (If a male rabbit with genotype GGbb crosses with a female rabbit with genotype ggBb, what will be the outcome?). Andria then asked, “what can we use to find out?” The students all answered in chorus “a Punnett square.” She called on a volunteer to create the Punnett square on the board and “when you finish can you please explain what you did?” Andria solicited students as resources for each other when she asked the student to explain their reasoning.

During class-work sessions, students worked on a task in groups or individually and Andria walked around and checked to see “where the students were” on the concept. She asked questions and guided their learning using context clues. Andria used systematic questioning, and she asked a question in response to student’s questions or answers. She responded to students’ questions with a reflective toss. One short example of this reflective discourse captured in Andria’s classroom was as follows

T what is going on here? (pointing at a picture)

S1 replication

T what is replication?

S1 making a copy

T what macromolecule replicates?

S2 DNA

T does protein replicates?

S2 no, protein…

Andria’s classroom discourse represented a complete ESRU (teacher Elicit, Student respond, teacher Recognize and Use) cycle. She echoed and restated students’ responses to questions so other students can process, and asked follow-up questions on the same concepts for students to explain.

Andria used an authentic task that required students to construct a response to continuously assess students' understanding of concept. Andria said, "I try to get something that is relatable and have an assessment for that thing the same day, so they do not forget" (I). For example, Andria gave students a task to recreate the processes of mitosis and meiosis and assessed students understanding by observing their actions, questions, and products. She also gave them a set of questions to assess their understanding and application of concepts. The focus questions that went with this task were

1) A cell is diploid and has seven pair of chromosomes. How many chromosomes would you expect to find in a haploid gamete?
2) A cell has 28 chromosomes in its gamete. How many chromosome pairs would you expect to find in the diploid cell that produced it?
3) A cell has 16 chromosomes in its normal 2N cell. How many chromosomes would you find in a gamete produced?

Students had to explain their reasoning and the method used. This example indicated how Andria used assessment as a tool for learning, to provide immediate feedback and feedforward into her instruction.

Andria also provided students with the opportunity for them to express their feelings, concerns, and questions about the lesson. She placed "sticky notes" on each table for students to "leave questions and ideas about the concepts" (O) on their way out at the end of class. This gave opportunity for students to reveal their thinking and concerns without fear of ridicule from peers. Andria also gave students ticket-in and ticket-out the door to check for understanding of past and present concepts.

Formative assessment practice: resources

Some resources that Andria used to help her with the practice of FA of students understanding of the core idea included, Google classroom, Socrative, and Pull Everywhere. She described that she administers the quiz on these resources and is graded "quickly and gives feedback to students instantly" (I). These resources also provided Andria with the flexibility of when to assess students. For example, during individual work time, Andria instructed two students who were absent during a quiz to take a "laptop and take your quiz." These technologies provided a "quick turnaround" and Andria acted on the feedback promptly. Case in point when Andria told one class, "yesterday we had an almost perfect score; many kids did well. So, we are going to move on from dihybrid cross." She used the missed questions for warmups for the next and subsequent day(s) or as a ticket out the door. Andria explained that she accesses the district's portal, *Edutrax,* for last year's questions and other online sites to pull questions that were related to the standard. She used sub-elements of these standards-based questions to create her warm-ups, to elicit students' prior knowledge and evidence of student learning or to know where the students are along the part towards the goal (mastery).

Formative assessment practice: challenges

The challenges implementing FA according to Andria is "finding questions that relate to what you want to teach, the time to grade them and … being able to analyze it critically and have the time to revisit the lesson" and fill in the gaps in students' learning. Andria believed that she could practice FA better, "if you have enough time you can research the questions and more, but time is of the essence." Another concern for Andria is that "there are always problems with connection." Andria used individual chrome books for students to take their Benchmark text for each unit and to take a quiz or makeup quiz and receive the results immediately. During one assessment episode, Andria gave students a quiz on monohybrid cross on the chrome books but was unable to obtain the response from all student because not all of them started on time. Some students needed help logging into the portal and some of the laptops were dead. These instances revealed some of the struggles Andria faced with implementing FA.

Three-Dimensional Teaching: Teacher's Conception of Integrated Science Instruction

When asked what Andria's understanding of 3D teaching was, she emphasized that "it is very difficult; this is our first-time doing 3D. We are still in the works of learning about 3D" (I). Andria's idea of teaching in this new way is to incorporate all the three dimensions at the same time, making sure her lessons reflect all the dimensions in the standard. She described that 3D teaching is a process where

> A teacher comes in with some phenomenon, and you see where the students are after brainstorming and discussing. Then wherever they are, based on the information you gather from your students, you start there to teach, or you lecture a little bit. Then you may give them hands-on activity for them to manipulate and try to see if they can figure out most of it. (I)

In this description of 3D teaching, Andria understood that a well-chosen phenomenon was necessary for the process's success. The phenomenon revealed what students knew, started and fueled conversations in the classroom, shaped the type of guidance provided to the students, and guided the subsequent activity. Andria believed that the ratio of teacher-talk to student-talk for 3D teaching should be 30:70. The teacher talks a little, to maximize the opportunity for students to learn by owning the conversation in the classroom. Andria explained that with this kind of teaching, she would probably "teach like 30% of the time and leave the rest" to the students. Adria stated that she comes in after the students "manipulate, sees if the students can answer the questions geared towards the lesson goal, if they can't figure it out, then I go back and lecture, and summarize for them." Andria believed that 3D learning requires students to figure out what is happening in the natural world using SEP and be able to apply the concept to another context using CCC to make connections. She emphasized that it is important for students to interact with all the dimensions to help them "solve problems that do not have answers to them yet."

Three-Dimensional Practice: Teacher's Enactment of Integrated Science Instruction

Andria started her lesson with careful planning. She explained that the preparation involved "making sure that I incorporate the DCI, CCC, and SEP" to address the standard. The dimensions were posted on one side of the board to inform the students of what they will be involved in for the day. A snapshot of the learning goal on her board looked like this

Standard	Obtain, evaluate, and communicate information to analyze how biological traits are passed on to successive generations.
DCI	Inheritance of biological traits
SEP	Asking questions and defining problems; developing and using models
CCC	Patterns (O)

When planning her lessons, Andria searched for many relatable phenomena guided by the core ideas to address different concepts. She provided students with a phenomenon usually a video, a picture, or a demonstration for the students to "describe and ask questions," and to reveal their thinking and what they know. She also asked students questions to "see where they are and clear up any misconceptions." One phenomenon that Andria said she used to introduce the concept of codominance was a picture of speckle chicken. She involved her students in SEP of making observations, inferences, and asking questions. Andria started with the phenomenon because "it is real life, relatable, gives students something to talk about and brings their attention to what we are trying to do for the day. It helps them think outside the box" (I). Andria used the engaged phase to sets the stage for her students, for them to start figuring out phenomena and solving problems

After the engaged phase, Andria gave them an "activity to explore." In the lesson observed, Andria selected the phenomenon and constructed the questions and activities from the standard containing the goal or learning target as shown on Table 6. below.

Table 6.
Andria's Practice of 3D

Lesson Title: Solving the mystery of green Parakeet parents with white, blue, and yellow offspring

Standard: Obtain, evaluate, and communicate information to analyze how biological traits are passed on to successive generations.

	Action	**DCI**	**SEP**	**CCC**
Pre-instruction	Plan from standard	Variation of traits	All related SEP	Patterns Cause & effects
Beginning	Warmup questions Phenomena (Green Parakeets with no green offspring) exercises Eliciting questions Class discussions	Variation of traits	Asking questions, Construct and use models,	Patterns
Middle	Constructing Punnett squares & determining ratios Group discussions	Variation of traits	Using mathematics & computational thinking Constructing explanations	Pattern Cause and effect
Last	Gathering evidence based on Mendel's principles of Dominance TOTD	Variation	Analyzing and interpreting data Constructing explanation and designing solution	Cause and effect

Andria also explained during the interview that the activities always mirrors the standard. The activity may ask students to use a "pattern… use a model, use cause and effect… doing ADI, or using a mathematical model for explanation."

Some of the activities were not standards-based, but Andria said she used them to provide her students with background information. Andria explained that "if they can manipulate this (auxiliary activity), then when they go to the explore part, they can manipulate that as well" (I).

Andria is referring to instances where she provided students with situations to define problems and ask questions, then use patterns or cause and effects to relate core ideas of the lesson to different contexts. For example, to guide students' learning, Andria gave them mnemonics and vocabulary tasks as auxiliary activities to facilitate students' exploration of concept and construction of foundational knowledge. She directed students' attention to these secondary activities, when she told them, "I need you to think about it," to help them "relate" to the big idea. In this way, Andria helped her students make connections between lesson's concept and the real word scenario.

Three-dimensional practice: resources

Andria relied on resources to help her with the implementation of 3D. She was able to "find some manipulatives that were easy to make and affordable for teachers and relatable to students" at the Stem-scope website. This source had manipulatives for the different stages of the 5E lesson (engage, explore, explain, expand, and evaluate) ready to use. (Andria said she incorporated elements of the 5E model into her practice of 3D to help her approach integrated instruction in a meaningful way and foster student learning). For example, to nurture students' developing knowledge of cell division, she gave them a blank printout of the phases of mitosis and meiosis (explore stage), cut out into individual stages for students to arrange in order and explain what is happening.

Three-dimensional practice: challenges

The challenges that Andria faced practicing 3D was "getting the materials" necessary to meet the different needs of her students and teach the standard. She explained that "trying to get things that the students understand or are familiar to them" is not easy. The common assumption is that we live in a technology era and information is readily available, however, "there is a lot of things we think students know, but they don't. So, when you use a picture, then you must explain what it is, that defeats many of the purposes." Although Andria indicated that "finding phenomena is not that difficult," she still must level the playing field by updating the background information for most of her students. For example, in one of her engagement activities, she used speckle chicken to elicit students' status, but she realized that her students did not know what that was, had "never seen one before." When Andria projected the picture, instead of focusing on the different colors on the chicken, the student asked whether the "chicken lays black and white eggs." Andria expected to generate a discussion on gene expression, in this case codominant, but because her students had not had the experience related to the background knowledge, she had to deviate from her intended target. The discussions according to Andria, did not advance students' understanding of the concept of inheritance. She had to explain the concepts to them instead of the students figuring it out for themselves or self-discovering.

Another challenge practicing 3D is "putting everything together, do a hands-on where they will look at stuff and be able to answer questions, but we don't have time for that one." For example, when students were demonstrating their skills of constructing Punnett squares on the board, the warning bell rung, and Andria said "oh almost time. If you didn't finish, I need it tomorrow. Finish it for homework." The beginning activities (warmups and the phenomena) took more than half of the class time, and Andria always carried over the remaining task to the following day. In anticipation, Andria assigned unfinished work or ticket out the door for homework and reminded students to visit google classroom for more practice.

Andria wished to have resources that provided her students with firsthand experience. She does not have them at present, and she expressed that "I want them to be able to look at a picture, but when you print it out, it does not look the same, so we do not have resources for that" (I). She envied some schools that have computers for each student and 3-d printers and believed it would be easier to teach 3D there.

Andria believed that collaboration would make a big difference, but 'we don't collaborate.' Some of the teachers "refuse to sit and collaborate." Collaboration is necessary for teachers according to Andria because it "helps us put everything necessary in our lessons and even our assessments" to make sure they address the three dimensions. Planning the lessons together, therefore, will assist the teacher's efforts towards 3D teaching because it helps to "bring different ideas together, different ways of assessing and different versions to meet different needs," so that the outcomes are a more standards-based lesson. Andria said "I will love to see more 3D teaching in action. I would like to observe someone teaching 3D. I want someone to do it so that I can see. Then I can do the same thing with my students." Andria believed that having an exemplar of 3D teaching is more than a necessity given that she teaches in a title one school where students are not as exposed to the real world. She wants to be certain that she is teaching the correct way or "heading towards the right direction."

FA3DL: Gaging Students' Integrated Science Knowledge and Adjusting 3D Instruction

Andria's Formative Assessment of Three-Dimensional Learning (FA3DL) starts with brainstorming ways to assess her students as she planned her lessons from the beginning using performance expectations of the standard. She started her planning with preparing her assessment, "making my ticket out the door as my formative, to gear towards the lesson's standard" (I). She inserted questions at strategic intervals in her lesson and prepared a ticket out the door for the end of her lesson, to constantly check for understanding. Andria used the DCI as the main organizing concept to construct learning claims for her lessons. Andria said she usually creates "at least five major questions from the core ideas that will help me know that they understand the lesson. I use them as the ticket out the door" (I, D). The questions usually prompt students to write what they understand or do not understand. Andria used questions that students missed as warm-up questions to help get them started on the next lesson. For difficult concepts where students still struggled, she used each "big question and make five mini questions" for use as a "quick quiz" (I), to assess specific skills. Here, Andria used the core ideas as tools to understand more complex concepts.

Andria assessed students' understanding of 3D learning using tasks that were standard-based and multi-component. An example of such assessment was a picture of green parent birds with offspring that were blue and yellow in color, but none of which were green. Andria used this phenomenon to assess the questions and claims generated by the students, how they construct their responses, and whether students could identify and used the appropriate CCC to investigate the problem. The students asked questions such as: why are the chicks blue and yellow? Are the chicks from the parent birds? Some of the students claimed that the chicks were from the parents but underwent mutation and others claimed that the chicks were from different parents. Andria redirected her students towards the previous unit and to use their knowledge of monohybrid and dihybrid crosses as guide to answer their questions. The students made models to investigate their claim (SEP) and analyzed the patterns (CCC) to explain the possibilities of genetic inheritance (DCI). The activity that followed this phenomenon assessed students' explanation to the questions, "what Mendelian Law made it possible? Where did the blue and yellow colors come from?" Andria explained that based on the standard, some of the tasks will "require students to explain their reasoning, give a scenario for them to ask questions, make a model, or analyze graphs" (I). She facilitated students practice with 3D learning by using "everyday activities and words" that replicate real world challenges to make connections with the dimensions.

Andria also designed tasks around clusters of related performance expectations that assessed how students used graphs, analyzed patterns, used mathematical reasoning, created models, and how students communicated information through presentations. For example, Andria involved her students in tasks that required students to complete Punnett squares to determine possible offspring and used mathematical reasoning to calculate the possible ratios and percentages of children. She also used an authentic task for students to apply standard-based knowledge and skills to real life situations using common vocabulary. She asked the students, "to grow your beautiful black hair, not some other person's ugly hair, what process will be possible, and why?" to facilitate student understanding of the concept. In this case, Andria related core ideas to student's interest and life experiences.

Andria checked for students' understanding during the manipulation stage of the activity, in small groups and on an individual basis. She approached each group and asked questions to check for understanding and provided help "on a smaller scale rather than lecturing." As Andria walked around from one table to another, she observed and listened for "specific skills and performances related to the standard" and made an instructional decision. For example, after she observed three groups modeling the process of mitosis, she called students attention and asked, "if a cell has to divide to produce two identical offspring, what must it do first and why?" She focused students' attentions to their models, the explanation they constructed, and to the connections they made.

Three-dimensional formative assessment in Andria's classroom can be summed up with the following vignette

T Somebody read the question

S In horses, color black B is dominant over chestnut b and gait trotting T is dominant over pacing t. Cross a homozygous black pacer with a chestnut heterozygous trotter.

T Stop. What comes to mind when you read this question? What can you use to help you answer this question? (identifying problem and proposing solution (SEP)

S Dihybrid cross

T Why is it dihybrid cross?

S It involves more than one

T More than one what?

S More than one trait

T Traits good. Can you identify these traits?

S First is color; the second is gait; (Ss- black is dominant, chestnut is recessive, Trotting is dominant, the pacing is recessive)

T Setup your dihybrid crosses (quick checked, students struggled constructing models). Can someone set up the dihybrid cross on the board? What is the parent genotype?

S I will

T Could you do me a favor and put the genotype of one of the parents to show them how you did it? Draw arrows so they would be able to follow how you got it. Can someone else help him? When I call on you, you complete a box.

T What do you guy notice? (looking at patterns (CCC))

S It repeats itself, the same (bell rung)

T What is the chance of having black trotter? Finish it and turn in as ticket in the door.

This assessment task explored how students used CCC (patterns) to connect SEP (identifying problems, asking questions and proposing solutions, and constructing models) to DCI (genetic inheritance).

After an assessment, Andria analyzed the results using excel to identify which questions students missed the most. These missed questions were "used as warmups, and I ask before we go over, what they were thinking… next time I will give them the same questions rephrased differently to see whether they understood it" (I). Missed questions were also "used as ticket-out the door in different versions." Sometimes these questions were assigned as homework for students to research about and bring the results in the next day as "ticker-in the door." Ticket-out assessed information that was just presented and was completed and turned in before exiting the classroom. While ticket-in assessed information the students did not finish in class and that needed more time to research and was turned in the next day as students entered the classroom.

Formative assessment of 3D: challenges

When asked what Andria's challenges were assessing 3D learning, she said "it is challenging to assess because it is hard to get all the dimensions into the test. I guess trying to word it in such a way that it gears towards the standard" (I), is difficult. Andria tried to include every dimension in the test, but it is hard to do that and "ask a question without giving away the answer." According to Andria, "with 3D assessment, you must beat around the bush for them to answer the question." This was Andria's first year trying 3D teaching and "standards-based questioning." She struggled looking for similar questions online, there are "not as many as the other ones, like asking *what cellular respiration is*?" Despite these challenges, Andria believed that 3D teaching is a "good model." Given that this is a new way of teaching, Andria is not confident about her understanding of this model and is not sure if she is practicing formative assessment accordingly.

Andria's Summary

FA according to Andria is a quick assessment given to students to see what the students know, to make sure they are learning, and provide evidence to make an instructional decision. Andria believed that FA gives a snapshot of where the students are on their learning journey. She explained that this evidence could be gathered using different means including warmups, TOTD, thumbs up or down, asking questions during discussions, and using sticky notes for students to leave comments. In her practice of FA, Andria started the lesson with tasks that elicited students' status. She checked or collected the warmups before reviewing them in class to see what students knew to help guide the next step of instruction and learning. The tasks were standard-based and authentic to assess skills and performances developed to solve real world problem as students learn. She used mostly questioning techniques during classroom discussions to assess students developing understanding of concepts. Andria used technology to check where her students were, for a "quick turn-around" and used the result the next day to adjust her lesson; however, internet connection was not always reliable.

Andria's conception of 3D teaching was that it is a difficult process. It involved students in interactions with all three dimensions in helping them solve new problems. Andria believed that 3D instruction started with choosing thoughtful phenomena to engage the students and move the conversation in the classroom forward. She explained that the students should do most of the talking and she comes in to fill the gaps. The ratio of such talk should be 70:30 in favor of the student, who must figure out the natural world for themselves.

She started her practice of 3D with planning her lessons to reflect the standard, which she shared with her students. She introduced the concepts by exposing her students to relatable standard-based phenomena that captured students' attention, involved them in some practices of the dimensions, and moved the discourse in the classroom forward. Following the phenomena, Andria involved her students in activities selected from the specification of the standard for them to interact with the three dimensions. Andria used available online resources to pull activities that address some aspects of the dimension for students' work. However, her primary challenge in practicing 3D teaching was assembling the necessary materials to meet the requirements of the standard and the needs of her students.

Andria assessed her students' three dimensional or integrated science learning with help from the standard. She planned her assessment tasks from elements of each dimension included in the standard alongside her activities. The questions were multi-components, assessed more than one performance expectations, and authentic. She used a big question (with all three dimensions) and divided it into small questions to assess specific performances. For example, Andria used a question for warmup to assess students' overall status and used parts of the same question to assess student learning of specific concepts during their manipulation activity. She analyzed the evidence gathered from these formative activities and used them to adjust students' learning and her instruction in different ways. Andria's challenge of FA3D learning was constructing tasks with all three dimensions included.

Cross Case Synthesis

Cross-case synthesis according to Yin (2009), is a method to aggregate findings across more than one case, through exploring, validating, and testing relations between concepts. The goal of this cross-case analysis was to accumulate case knowledge about teachers' understanding and practice of FA and 3D teaching, and to facilitate the comparison of similar and contrasting points for a robust finding. A word table was created to display common patterns based on identified uniform categories that emerged from the different teachers.

Teachers' Understanding of Formative Assessment

Table 7.
Teachers' Understanding of FA

Category	Chelsey	Paul	Andria
Description	Checkpoint	Constant checks	Snapshots
What is it	Anything to elicit status & gauge learning	Anything to elicit status & gauge learning	Anything to elicit status & gauge learning
When	Before, during, & after instruction	Before, during, & after instruction	Before, during, & after instruction
Who	Teacher and student	Teacher and student	Teacher and student
Purpose	Provide evidence to adjust learning and instruction	Provide evidence to adjust learning and instruction	Provide evidence to adjust learning and instruction
How	Test, quiz-, listening, observation, questioning, discussion	Test, quiz, bell ringer, TOTD, argument, discussion, observation, listening, questioning,	Quiz, TOTD, thumps up or down, sticky notes, questioning observation listening, discussion,
Impact	Reinforces concept	Brings clarity	Reinforces concept

Ticket out the door- TOTD

A visual synopsis of Chelsey's, Paul's, and Andria's understanding of FA is displayed on Figure 7 above. All three teachers identified FA with a different name, however, their descriptions of FA or what it does were similar. Chelsey called FA a "checkpoint" for understanding and described it as a process that informed her of where her students were, that gaged understanding of student learning, and that provided her with background knowledge of where to go next. Paul called FA a "constant check of what we do" and described it as a process that continuously informed him of how well his students have learned the topic, of where they were, and whether to remediate or not. Andria called FA a "snapshot" of the students and described it as a process that immediately informed her where her children were, whether they understood the lesson, and gave her evidence to either move forward or reteach. These three teachers understood that FA was a tool to (1) elicit their students' status (where the students are initially) (2) measure progress (what the students have learned) (3) provide evidence for instructional decision (take the next step). Chelsey, Paul, and Andria similarly explained that they could use anything (for example, the hand, a quiz, TOTD, and questioning) to accomplish these three tasks. Chelsey and Andria shared that FA reinforced the concepts for students, while Paul believed that it brought clarity to both teaching and learning. Formative assessment according to the trio, was necessary to help the teacher, and the student grow. Contrary to Chelsey and Andria, Paul in his description of FA recognized the role of students in the formative process.

How Teachers Track What Students Know, Have Learned, and Adjust Science Instruction

Table 8.
Teachers' Practice of FA

Category	Chelsey	Paul	Andria
Success criteria	Identified from standard and shared	Identified from standard and shared	Identified from standard and shared
Elicit status	Used quiz, phenomena, and Socratic questioning	Used warmup, quiz, phenomena, Socratic questioning, rough drafts	Use warmup, quiz, Socratic questioning, phenomenon
Tasks to guide and gage learning	Purposeful FA, created from standard TOTD	Purposeful FA, created from standard TOTD	Purposeful FA, created from standard TOTD, TITD
Process	Observed, listened, used reflective toss, and evaluated students' skills and performances using SB-rubric	Observed, listened, used reflective toss, and evaluated students' skills and performances guided by the standard	Observed, listened, used reflective toss, and evaluated students' skills and performances guided by the standard
Student role	Solicited students as masters of their own learning Fostered self and peer evaluation	Solicited students as masters of their own learning Fostered self and peer evaluation	Solicited students as masters of their own learning Fostered self and peer evaluation
Next step	Self-reflection, error analysis, review, revisit concept, google classroom	Error analysis, tutoring, revisit concepts, warmups, google classroom	Error analysis, tutoring, warmups, revisit concept, google classroom

TITD- Ticket in the door; TOTD- Ticket Out The door

A cross-case synthesis of how Chelsey, Paul, and Andria tracked what students knew, what students had learned, and how they adjusted their science instruction and close the gap in student learning indicated more similarities than differences as shown on Table 8. above. All three teachers started their lessons by identifying and sharing the goal of the lesson (standard and learning targets) or the success criteria. Chelsey always gave the "expectation for the day," Paul always made it "clear where the students were heading," and Andria informed them of what they would "be able to do at the end." These goals were posted at a corner on the board for all three teachers, to remind students daily of where they were going. Though, Paul and Andria were more consistent than Chelsey in reminding students of the daily objectives.

To elicit students' status or where students were, Paul and Andria used warmups daily. Paul had his warmup written below the learning target ready for students to start on as they walked in. Andria had her warmups projected on the monitor ready for students to start on as they walked in as well. Chelsey was not regular with her warmups; she gave a warmup once when she had to step out at the beginning of class. Paul and Andria encouraged their students to complete the warmups individually before the whole class reviewed them to see where students were. All three teachers used a thought stimulating technique, Socratic questioning to solicit multiple responses from students and to assess their evolving understanding of concepts. Students' questions were also answered similarly with a reflective toss by all three teachers.

During the discussions that came after the warmups or phenomena that preceded student activities, all the teachers mostly answered students' questions with another question that assessed students evolving understanding of concepts. This reflective questioning also gave students the opportunity to answer their questions. For example, Paul brainstormed how he would "get the students to answer this question themselves." They all encouraged their students to ask questions and answer the questions themselves and to assemble resources to help them think through and solve problems. They all similarly fostered self and peer evaluation, thus solicited students to be resources for each other and master of their learning.

The assessment tasks selected by all three teachers were purposeful and created with the standard in mind, to assess students learning of specific concepts. For example, Chelsey was deliberate in her assessment as she said that, "any task that I give, there is something specific I want them to learn from it." In the same light, Andria addressed the learning goal by creating an "assessment for that concept at the same time" as the activity. Also, Paul said that how he "assess is based on the concept" he is teaching. All three teachers similarly used assessment tasks that reflected the dimensions in the standard and they, therefore, assessed what the students had learned about the standard. The trio also challenged their students to be self-directed and to be resources for each other. In one case, Paul explained to his students that "my knowledge about the topic can't be the only resource," and encouraged them to "bounce your ideas off each other."

Chelsey, Paul, and Andria acted similarly on evidence collected to adjust their lesson and students' learning or to take the next step of instruction. They created remediation activities for students to complete on Google classroom and revisited difficult concepts in their subsequent lessons. Paul and Andria used difficult concepts as their warmups for continuous reinforcement of specific ideas. Chelsey used the information in future instructions and reassessed them. All three teachers involved their students in similar remediation activities including error analysis, tutoring, and revision. Chelsey also encouraged her students to complete individual reflections, and Andria encouraged her students to leave comments on sticky notes. Information from these activities was also used to adjust instruction. For example, Andria changed her lesson of the day for one of her classes based on the feedback she got from a quiz. She told the students, "you guys did good in the quiz, so we are going to move straight into dihybrid crosses." Paul and Andria were also more consistent in taking the next step in instruction as they used missed questions for their warmups daily. Their next step also influenced their future lessons.

Chelsey, Paul, and Andria all shared during the interview that using technology like Poll everywhere and Kahoot helped with implementation. It provided immediate results about students' status and what they had learned to guide the next step of instruction. However, they encountered some problems from occasional glitches with computers and with limited time to assess students with variable abilities.

Teachers' Conception of Integrated Science Instruction

Table 9.
Teachers' Understanding of 3D

Category	Chelsey	Paul	Andria
Degree of complexity	Complex	Complex	Complex
Time to practice	Long	Long	Long
Integrate dimensions	Can integrate all three dimensions, but not in a day	Can integrate small pieces of each dimensions daily	Can integrate small pieces of each dimensions daily
Process driven by	Phenomena Relatable and nonstop engaging Wonderment	Phenomena Relatable and nonstop engaging Rationale for learning	Phenomena Relatable and nonstop engaging Thinking out the box
Purpose	To solve a problem without an answer yet	To solve a problem without an answer yet	To solve a problem without an answer yet
Classroom talk	S-talk > T-talk	S-talk (65) > T-talk (35	S-talk (70) > T-talk (30)
Learning environment	Student centered	Student centered	Student centered

S- student; T- teacher

Chelsey, Paul, and Adria all expressed little confidence about their understanding of the process of 3D teaching. All three teachers believed that it was a complex process. Table 9 summaries understanding of 3D teaching from Chelsey, Paul, and Andria. When asked about 3D teaching, Chelsey said that it was "complicated," Paul said that it was "tricky," and Andria said that it was "arduous." Chelsey described the process as related to "teaching all three dimensions together and not separately as it has always been done" in the past. Paul understood that it was "important to teach all those concepts in the curriculum that supports the standard, that is the way it should be every day." Andria described the process as related to putting "everything together, all the dimensions to address the standard." To practice 3D teaching, everything (all three dimensions) must be included.

All three teachers indicated that the process started with some phenomena that are relatable to students' interest and experiences, to engage students from beginning to end. Paul believed that it provided a rationale for students to learn the concepts and Andria believed that it helped the students think outside the box. The three teachers also indicated that it took time to implement 3D teaching. While Chelsey thinks she cannot teach the three dimensions in one day. Paul and Andria conversely believed they could by incorporating small elements of each dimension into the lesson. These teachers had similar beliefs about the ratio of teacher talk versus student talk in the classroom. All three teachers equally thought that the nature of discourse in the classroom should favor student-talk over teacher-talk. Therefore, providing the context (interest and experience related phenomena) for such discourse to occur is paramount to 3D teaching. They all believed that these classroom conversations afforded the students opportunity to figure out the phenomena or the natural world. Chelsey, Paul, and Andria all believed that the purpose of 3D teaching was to prepare students to be able to "solve problems that do not have answers yet."

Teachers' Enactment of Integrated Science Instruction

Table 10.
Teachers' Practice of 3D

	Teacher	Action	DCI	SEP	CCC
Pre-instruction	**Chelsey** **Paul** **Andria**	Plan from standard	Inheritance Variation of traits	All related SEP	Patterns Cause & effect
Beginning	**Chelsey**	Reviewing concepts Phenomena (genetic disorder-Hemophilia) Class discussions		Asking questions (making claims and proposing solutions)	
	Paul	Warmup questions Phenomena (genetic disorder-Huntington) Class discussions	Inheritance		Pattern Cause & effect
	Andria	Warmup questions, TITD Phenomena (Green Parakeets parents with no green offspring) exercises Eliciting questions Class discussions	Variation of traits	Asking questions, Construct and use models,	Patterns
Middle	**Chelsey**	Identifying and researching on a genetic disorder (individual/group) Group discussions	Inheritance	Planning and carrying out investigations	Not observed
	Paul	Constructing pedigrees and determining outcomes Group discussions		Analyzing and interpreting data Constructing explanations	
	Andria	Constructing Punnett squares & determining ratios Group discussions	Variation of traits	Using mathematics & computational thinking Constructing explanations	Pattern Cause and effect
Last	**Chelsey**	Continuing research and preparing presentation Homework			Cause & effect
	Paul	Gathering evidence based on Mendel's law of dominance to support results, TOTD Homework- identify disorder & use pedigree to explain inheritance in google classroom	Inheritance	Obtain, evaluate, and communicate information	Pattern Cause and effect
	Andria	Gathering evidence based on Mendel's principles of Dominance TOTD	Variation	Analyzing and interpreting data Constructing explanation and designing solution	Cause and effect

TOTD- Ticket Out The door; TITD- Ticket In The Door

An image of how the teachers practiced 3D teaching can be formed from Table 10 above. A cross-case analysis of Chelsey's, Paul's, and Andria's enactment of 3D teaching indicated that all three teachers started with careful planning to make sure they incorporated all three dimensions of the standard into their lessons. The teachers each introduced the concept with a phenomenon, carefully selected to meet the performance expectations and relatable to students' interest and experiences. Chelsey had no routine for the start of her lesson. Sometimes she started with a review of the previous day's lesson, the students continued where they left off, or she started with a phenomenon. Conversely, Paul and Andria always started their lesson with a warmup followed with either phenomena or mini activity.

The teachers used relatable phenomena to serve as fuel to ignite student's curiosity and start the conversations in the classroom. While Chelsey and Paul focused on using true stories, Andria on the contrary used what was interesting to the students. For example, Chelsey and Paul used an episode of Maury and of Snap as their phenomenon. Classroom observations indicated that these phenomena challenged students thinking and triggered wonderment, as they engaged in SEP of asking questions, making claims and proposing ways to investigate the problems to gather evidence, and participating in argumentation from evidence.

Chelsey, Paul and Andria also used the beginning activities to involve the students in using the CCC, mostly cause and effect and patterns to make connections between the SEP and the DCI in their discussions. The discussions in the classroom were usually interesting and ranged about 2 - 20 minutes. In all three classrooms, the discussions were dialogic and the investigations where student led. The students worked in groups and relied on each other for help rather than the teacher. However, Chelsey's classroom environment was more student centered, she provided the least guidance and her students were more self-directed. All three teachers recognized that it took time to practice the three dimensions together. Therefore, they all exercised some flexibility in implementing 3D teaching.

Similar patterns were observed in activities that usually followed the discussions from the phenomena, they were all open-ended. The teachers provided minimal guidance in the beginning and left the decision on how to approach the activities to the students. One culminating activity that all three teachers involved their students in were ADI investigations. Likewise, all three teachers implemented one ADI investigation entitled "Are all of Mr. Johnson's children his biological offspring?" In this activity, the students participated in the practices of designing and conducting investigations, analyzing data, and arguing from evidence. The students also used a relevant CCC, cause and effect, to investigate the reaction of blood with different antigens when in contact with an indicator and another CCC, patterns, to investigate the possibility of inheritance of a given blood group using Punnett squares.

The three teachers also utilized auxiliary activities that may or may not address the standard, but the goal was to prepare the students for the main task (ADI) that reflected the elements of the three dimensions. They each used a model that they were more comfortable with to facilitate their implementation of 3D. Chelsey utilized case studies, Paul the CER activities, and Andria the manipulatives (5E format) as their supporting activities to gradually develop students' performance and skills towards the goal of the lesson. Of the three teachers, Paul was most oriented towards 3D instruction (he used more phenomena and involved his students more in the SEP and CCC) and Chelsey was least oriented towards 3D instruction.

How Teachers Gaged Students' Integrated Science Knowledge and Skills and Adjusted Instruction

Chelsey, Paul, and Andria all planned their assessment tasks alongside the performance activities, taking into consideration the elements of the different dimensions specified in the standard. The assessments tasks were thus purposefully designed to include multiple components, variable forms, and attentive to the different dimensions. However, the assessment tasks were implemented piece meal to assess students evolving understanding. For example, the integrated assessment tasks were implemented to measure each dimension at a time in the beginning, so at the end, students could answer the big question. All three teachers wanted to assess how students obtained, evaluated, and communicated information to analyze how biological traits are passed down to successive generations. All three teachers similarly chose phenomena based on the core ideas of the standard. Therefore, the goal was to assess the SEP students used to figure out the phenomena or the CCC used to connect the SEP to the DCI. For instance, the kinds of questions students asked, the claims they advanced, or how they proposed to conduct their investigation and to solve the problem. The goal was also to assess how students used the evidence collected to support their claims and how they used patterns to make connections. For example, Chelsey Paul, and Andria assessed the types of model students used to investigate the possibility of paternity and how students analyzed the models for patterns of inheritance.

Table 11.
Teachers' Frequency of Planned and Unplanned FA of 3D Leaning

	Formative Assessment of Three-Dimensional Learning					
	Unplanned- (discussions/questions)			Planned- (lesson plan/quiz)		
	DCI	SEP	CCC	DCI	SEP	CCC
Chelsey	Most of the time	Sometimes	Sometimes	Every time	Sometimes	Never
Paul	Every time	Most of the time	Most of the time	Every time	Sometimes	Sometimes
Andria	Every time	Most of the time	Sometimes	Every time	Sometimes	Sometime

Table 11. above shows the extend of teacher's unplanned and planned formative assessment of three-dimensional learning. It was observed that Paul was more consistent in using all three dimensions than Andria and Chelsey. It was also observed that Paul and Andria utilized more planned FA of 3D than Chelsey. However, during the interview, Chelsey articulated a better understanding of FA than Paul and Andria, but not in practice. During planned FA, Chelsey, Paul and Andria utilized the DCI every time (one hundred percent as the focus with all queries). Whereas during the unplanned FA, Paul and Andria used the DCI in all observations but Chelsey used it most of the time (about seventy five percent as the focus with all queries). Another difference was observed during the planned and unplanned FA with the teachers' used of CCC both. For the planned FA, Paul and Andria assessed CCC sometimes, but Chelsey never assessed students' understanding of CCC. While during the unplanned FA, Paul used the CCC most of the time, but Chelsey and Andria used it sometimes (about twenty five percent of the time). It was observed that with planned FA the three teachers embedded the DCI in all their assessment task, but the SEP and CCC were absent except for one task, and again Paul utilized the CCCs more.

Table 12.
Teachers FA3D Learning

	Chelsey		Paul		Andria	
FA Element	**Tasks Type**	**Dimension**	**Tasks Type**	**Dimension**	**Tasks Type**	**Dimension**
Planning Where is S going	Success criteria Standard	CI: Inheritance SEP: assess all practices CCC: patterns, cause & effect	Success Criteria Standard	CI: Inheritance SEP: assess all practices CCC: patterns, cause & effects	Success Criteria Standard	CI: Inheritance SEP: assess all practices CCC: patterns, cause & effect
Elicit Where is S now	Phenomena ChR., CR	CI: Hereditary SEP: Asking questions; Solving problems Developing & using models Constructing explanation Engaging in arguments from evidence CCC: patterns, cause & effect	Warmup Phenomena ChR., CR, SR	CI: Hereditary SEP: Asking questions; Solving problems; Developing & using models Constructing explanation Engaging in arguments from evidence CCC: patterns, cause & effect	Warmup Phenomena, TITD ChR., CR, SR	CI: Hereditary SEP: Asking questions; Solving problems; Developing & using models Constructing explanation Engaging in arguments from evidence CCC: patterns, cause & effect
Navigation How to get there	Quiz, TOTD Mini task CR, SR, ChR Observation Listening	CI: Hereditary SEP: assess all practices CCC: patterns, cause & effect	Quiz, rough drafts, TOTD Mini tasks/ CR, SR, ChR Observation Listening	CI: Hereditary SEP: assess all practices CCC: patterns, cause & effect	Quiz, TOTD Mini tasks/ CR, SR, ChR Observation Listening	CI: Hereditary SEP: assess all practices CCC: patterns, cause & effect
Next step Adjustment	Error analysis, reflections, CR	CI: Hereditary SEP: Based on deficiency CCC: patterns, cause & effect	Error analysis, re-quiz CR, SR	CI: Hereditary SEP: Based on Deficiency CCC: patterns, cause & effect	Error analysis, re-quiz CR, SR	CI: Hereditary SEP: Based on deficiency CCC: patterns, cause & effect

Student- S; ChR- chorus response; SR- selective response; CR- constructive response; TITD- Ticket In The Door; TOTD- Ticket Out The Door

Table 12 above painted a detailed picture of the teachers' formative assessment of 3D learning. To assess student's status of the core idea and evolving understanding of concepts, Chelsey used tasks that solicited mostly chorus responses and constructive responses (open call questions, error analysis, and reflections). Whereas, Paul and Andria in addition to soliciting chorus and constructive responses, used selective responses to evaluate students understanding (using previous benchmark questions). During participation in classroom activities, all three teachers observed and listened to students and assessed the performances and the skills they used, to gauge what students had learned and what needed to be learned. All three teachers likewise encouraged students to practice self and peer evaluation, to take ownership of their learning.

Assessment is formative to the extent that the evidence of students' learning is used to adjust student learning and instruction. All three teachers used evidence from assessing students' performances and skills to make instructional decisions with minor differences. For this next step of instruction, the trio revisited challenging concepts in future lessons, as warmups or in discussions, provided tutoring, and provided opportunities for continuous practice of concept through Google classroom platform. Chelsey in addition to these steps, required her students to complete reflection activities on their strengths and weaknesses that may reveal the reasons they missed the questions and expose misconceptions. Paul and Andria in addition, used the missed concepts as warmup to uncover students thinking and clear up misconceptions. Alongside these strategies, Andria also assigned challenging concepts for homework for student to research and bring in their findings as ticket in the door (TITD).

5. DISCUSSION

The purpose of this study was to explore science teachers' understanding of formative assessments and how they elicited and interpreted students' assimilated science knowledge to adjust instruction and improve students' three-dimensional learning. The study also described teachers' understanding and practice of three-dimensional teaching from the perspectives of the teachers. Four exploratory questions guided this research

1. What are high school science teachers' understandings of formative assessment?
2. How do high school science teachers understand three-dimensional science instruction?
3. How do high school science teachers practice three-dimensional teaching?
4. How do high school science teachers practice formative assessment of three-dimensional learning?

In the first part of this chapter, findings from the cross-case synthesis for the three teachers were presented to answer the research questions and discussions of the study as it relates to the foundations of FA and 3D teaching and learning. The findings were positioned in relation to other findings from current literature. The second section of the chapter discusses the limitations, implications, and recommendations for future research.

Question 1.

What are High School Science Teachers' Understanding of Formative Assessment?

All three teachers shared a similar understanding of the concept of formative assessment. Chelsey's, Paul's, and Andria's understanding of FA were that it is anything that helped them gauge what students know, have learned, and helped them decide on what to do next. Formative assessment, as described by Black and Wiliam (1998), is "all activities undertaken by teachers and or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (pp. 7-8). The description of FA provided by the three teachers captured the what, how, where, and the why of Black and Wiliam's definition. The teachers indicated that FA is anything (quiz, hands, observation, questioning, and listening) (the What) used to elicit information (the How) about where the students are in their learning, to use the information to make instructional decision (the Why) and to close the gap or to meet the learning goal (the where). The teachers' descriptions of FA did not highlight the important role students play in the formative process (the Who, in this case, a partnership between the teacher and the students). This omission does not mean that the students were left out entirely because all three teachers believed that the lack of student participation was a stumbling block to the formative process. If this omission had been purposeful, then the teachers' description of FA would have been incomplete, given that the practice of FA gives voice and values students' experiences in discussions and development of knowledge (Ruiz-Primo & Furtak, 2007).

Formative assessment is valuable to the extent that it informs the teaching and learning. To clarify, this point, Chappuis (2009), argued that "it is not the instrument that is formative; it is the use of the information gathered, by whatever means, to adjust teaching and learning, that merits the "formative" label" (p. 4). All three teachers believed that the evidence from FA of students' learning should be used to take the next step in instruction, to remediate (for less difficult concepts), reteach the material (for complicated ideas), or challenge students thinking in future lessons. The teachers were confident that evidence from FA helped reinforce or brought clarity to the teaching and the learning occurring in the classroom. They understood the necessity of formative assessment outcomes to make a better decision to their instruction and student learning. Black and Wiliam (2009) support these teachers' beliefs with the argument that evidence is used to "make decisions about the next steps in instruction that are likely to be better, or better founded than the decision they would have taken in the absence of the evidence (p. 7).

Chelsey, Paul, and Andria's, explanation of FA (the What), involved an element of scrutiny. They used the words check-point (like the teachers in Phelan et al.'s, 2011 study), continuous-check, and snapshot, respectively, to reference moments when students' learning was inspected during instruction to take immediate action. This informal assessment, according to Ruiz-Primo and Furtak (2007), allows the teacher to react impromptu or on the fly to students' ideas throughout the lesson to move the learning forward. The three teachers also considered the moment to be a point at the end of instruction when they proposed using tests or TOTD. With formal FA, a teacher planned tasks in preparation for this moment. In both the formal and informal FA, it is timely as the teacher uses the information immediately to make instructional adjustments (remediate and reteach concept and skills) and help build initial abilities to sustain a shift in learning over time (Black & Wiliam, 1998; Sadler, 1989).

The teachers in this study had a general understanding of formative assessment to be as a descriptive, interpretive, and steering process. Formative assessment informs students of the goal or success criteria of the lesson, it gages where students are based on the goal, and it uses the evidence to adjust the teaching and learning (Black & Wiliam, 1998; Bloom, 1969). For a rounded understanding of FA, a teacher must first understand the goal of the lesson, identify tasks to elicit students' status, and what students will do to close the gap in their learning goal. Secondly, the teacher must share the lesson goal with the students, and both should accurately interpret the evidence to close the gap in instruction and student learning.

Science teachers' understanding of FA was summarized as any tool to elicit students' information (how they are reasoning and what they have learned) during instruction and at the end of instruction, and the evidence obtained is used to adjust teachers' instruction. These teachers did not highlight the process of collaboration with students, which is necessary for effective FA practice.

Question 2.

How do High School Science Teachers Understand Three-Dimensional Science Instruction?

All three teachers felt that 3D teaching was a challenging and daunting process. Chelsey, Paul, and Andria all expressed their feelings towards 3D instruction using words that reflected a degree of complexity like complicated, tricky, and very difficult, respectively. Their attitudes towards 3D teaching were not surprising, given that the teachers were in their first year of experiencing 3D teaching. There is much more information available on existing literature for two-dimensional teaching and learning (how science teachers merge DCI and SEP). However, research on three-dimensional (integrating DCI, SEP, and CCC) is mostly theoretical with minimal application to classroom practice or research (Fick, 2017; Fick & Songer, 2017; Harris et al., 2015). Therefore, these teachers did not have a model to follow from the start. Although they received training on how to implement the ADI investigations, and they utilized it as their primary activity to integrate the dimensions, they did not consider that as preparation for 3D teaching.

The teachers' descriptions of their understanding of 3D teaching were similar. They said that it involved teaching the three dimensions together, the Science and Engineering Practices (SEP), the Crosscutting Concepts (CCC), and Disciplinary Core Idea (DCI). Their thoughts were in line with the Framework for K-12 Science Education (NRC, 2012) and NGSS (NGSS Lead States, 2013) vision, which advocates for the integration of elements for 3D teaching and learning. The teachers explained that 3D teaching usually started with a phenomenon that was interesting and relatable to students' experiences or using something students were already familiar with to engage and help them figure out phenomena and solve problems (NGSS, 2016). The use of phenomena was necessary because it captured students' attention and engaged them throughout in progressive science thinking and helped illuminate connections and importance of science ideas, as stated in the National Academic of Science, Engineering, and Medicine (NASEM; 2017).

The most significant influence of science learning occurs when the teacher becomes the learner in their practice, and the student becomes the teacher of their knowledge (Hattie, 2012). Chelsey, Paul, and Andria believed that classroom talk is vital to learning and that the student should do most of the talking. Paul and Andria went as far as attaching a ratio to the nature of discussion in their classrooms; Paul said 65% for students vs. 35% for teacher and Andria said 70% for students vs. 30% for the teacher. These teachers believed in sharing the responsibility and power in their classrooms so that teaching and learning combine towards a shared goal for learning (Heritage, 2011). With students doing most of the talking, it reveals the different discourses that are occurring in the classroom and allows the teacher the opportunity to see or discover the patterns and underlying meanings +hidden in classroom talk (Lemke, 1990). All three teachers believed that the goal of three-dimensional teaching was to take students to a point where they can apply concepts, they learned to solve problems that do not have an answer yet. This thinking is parallel to the goal identified in the NGSS (NGSS, 2014). To accomplish this goal, the teachers believed it essential to change the way they had been doing things. This change for three-dimensional teaching, according to Krajcik (2015), "is an orientation one takes to science teaching" (p. 16). He suggested that teachers should focus on how students are making sense of phenomena or designing solutions to a problem rather than focusing on how often students used the three dimensions. For these teachers, it meant not giving students the answer but creating opportunities through merging the dimensions, for students to wrestle with their ideas and figure out for themselves how the world works. Taking this direction was necessary to get students to wonder and start figuring out stuff and applying their learning to different contexts as scientists do. Three-dimensional teaching suggests a shift in classroom roles for students from receptors to creators of knowledge (NRC, 2012; NGSS, 2014). The teachers in this study also believed that giving students a more significant role in their learning will create a classroom climate where students can ask questions, make claims, propose how to answer their questions about the phenomena through investigations, and by integrating the dimensions.

To conclude, three-dimensional teaching, according to the teachers in this study, is a challenging concept that integrates all three dimensions. This kind of instruction starts with introducing a phenomenon to capture students' attention, involving students in thinking, and moving the discussions in the classroom forward so that students can solve problems without an answer yet.

Question 3.

How do High School Science Teachers Practice Three-Dimensional Teaching?

Practicing 3D teaching, according to the National Research Council (NRC, 2012), is the process whereby the teacher integrates the DCIs, SEPs, and CCCs. They used the core ideas in the DCI to organize their instruction and their assessment and used concepts from different disciplines in CCC to facilitate students' understanding of phenomena or the natural world. The teachers involved their students in scientific behavior of SEPs, to help them explain or figure out phenomena. Chelsey, Paul, and Andria all planned their lessons with the core ideas of the standard to make sure they aligned their lesson goals with the goal of the standard. They chose, at the same time, the corresponding phenomena that were also relatable to students' interests and experiences. Following the phenomena, the teachers always provided the students with activities to involve them with the elements of the different dimensions. Based on the Framework's vision (NRC, 2012), there is some evidence that Chelsey, Paul, and Andria were moving towards adopting the reform message.

The teachers started their lessons with a phenomenon, usually true stories, something the students were familiar with, that engaged them with the concepts of the core ideas to generate discussions and to make connections in their explanations. It can also be a "puzzle or something counterintuitive" (NASEM), 2017, p. 16) to activate students' thinking. Besides, a phenomenon that has been given considerable thought, "will focus students on connections between what they are learning and what they observe in the world… provide students with a shared experience to which they all have equal access (p. 16). However, cultural diversity among students resulted in phenomena that did not provide shared experience to all students, and the teachers were challenged by it. The teachers in Reiser et al. (2017) study after professional development sessions and practice were able to improve their understanding of phenomena and development of models. Also, they improved their ability to structure explanations around CCC. The teachers in this study also used CCC as a lens to explore the phenomena and as part of students' conceptual models to illuminate their understanding (Fick, 2017), but to a limited degree. For example, one teacher used the case of people who are color blind but can drive to model how people use patterns in nature to solve real-life problems (using changing sequence of light signals to help with their driving). Another teacher stated that,

> During classroom discussions, the students were doing most of the talking. They were
> Asking questions- why are none of the offspring color like their parents?

Making claims- we believe that the difference in color is due to independent assortment

Suggesting ways to investigate problems- we can construct Punnett squares and calculate the probability to see how the trait was passed down (SEP).

The teachers' intentions were to avoid front-loading the information, but rather to provide bare-bone information and come in to fill the gaps in students' learning and summarize. However, the teachers also wanted to communicate the right information and joined the classroom conversation with questions and cues to orient the students thinking and controlled the flow of knowledge.

Following the discussions and visualizations from the phenomena, the teachers involved their students in activities like a case study, CER, 5Es, and ADI investigations, to interact with the different elements of the three dimensions. The authors of ADI recognized the importance of scientific argumentation in science education and the science classrooms (Grooms, Enderle, & Sampson, 2015). Involving students in scientific argumentation is essential to help them develop and enhance scientific knowledge. Students thus are engaged with different science practices to boost their understanding of concepts. During the argumentative inquiry, Chelsey, Paul, and Andria involved their students in the SEP (asking questions, making claims, reasoning from evidence, generating and interpreting data, participating in social argumentation sessions, and reviewing each other's work and writing their investigation reports). They also involved students in using cause and effects (CCC) to determine reactions to certain chemicals and in using structure and function to analyze the reactions of different blood groups to antibodies based on the antigens attached on their surfaces. Andria designed instructional sequence using the 5E model to guide students through the critical steps of building new knowledge for three-dimensional learning. Drawing from the constructionists' approach, students can thus construct knowledge of the world through their individual experiences (Bybee, 2015). As stated in the literature, the 5E model can ease the problem of merging the three dimensions in the science classroom and activities selected for each phase to give students the opportunities to experience SEP, CCC, and DCI (Bybee, 2013).

Students in these teachers' classrooms completed their work in groups. These teachers always assembled and reassembled students into a grouping of varying sizes that "makes for sound learning" (Hock, 1961, p.421). For example, during ADIs or other culminating activities, these teachers entrusted students with the tasks of appreciating their strength and weaknesses and guiding their peers, to bounce their ideas off each other (S1- Is this a good question? S2 to S3- Listen to my evidence, S3 to S2- That is data, not evidence); to use each other as resources (where you' all find that at? Look on the back, at the bottom (group1 (S1, S2, S3, S4) to group2 (S1, S2, S3, S4) support)); to assess each other's performance by provide constructive feedback (you have the graph, nice. Add Punnett square to show parents). The importance of grouping students is reflective of an effective classroom and according to Hock (1961),

> We group to provide for the vast differences that exist among any aggregation of individuals. The great varieties of interest and purposes, the wide range of talents and skills, the important differences in ability and potential, in speed, depth, and nature of comprehension… that provide opportunity for each student to move rapidly as possible in reaching his own potential (p. 421).

Chelsey, Paul, and Andria used grouping in their classrooms to leverage differences in their students so together they can address the varying scope of information, understanding, and attitudes.

Studies suggest that CCC can serve as a lens to analyze phenomena, as a bridge to connect concepts across disciplines, as a tool to understand the content in different ways depending on the purpose and nature of the inquiry, and as a rule to guide the use of DCI (Rivet et al., 2017). Although the teachers used CCCs to analyze the phenomena and frame their classroom discussion activities, they did not exploit its full potential. They used mostly cause and effect and patterns (the two that connects the DCIs of the standard to the SEPs). They sometimes used structure and function (Teacher- why does a harmer have two different heads? Student- because it has two different functions), stability and change (Student- If they are identical twins, why are they different? Teacher- what guarantees that everyone is different?), and energy and matter (T- matter cannot be created nor destroyed) as preparation to get students in the habit of using CCCs. However, this practice was not consistent.

Exercising flexibility is necessary for the practice of 3D teaching (NRC, 2012). Chelsey, Paul, and Andria demonstrated flexibility in their implementation of 3D so students can explore concepts freely (NASEM, 2017. Although all three teachers taught Biology, a subject with an associated high stakes test, they were not restrained by the clock but allowed the conversations in the classrooms to guide the path they took. The students could approach the activities the way they wanted to, to find solutions to problems. Also, the opening activities (warmups and phenomena) usually extended into the middle of the lesson and determined how much time the subsequent tasks had available. However, the teachers were not too concerned with the discussions taking up much of the class time because the students were able to explore the concepts and answer most of the questions in the main activities during this time. For instance, one of the teachers told the students, "look at your learning target for today, you have already covered half of it, and you have not even started your work session yet."

Research informs us that teaching with classroom discussion stimulates critical thinking. Classroom discussion challenges students to think deeply and communicate ideas clearly and provide them with opportunities to ask and answer questions to assess their learning (Brookfield & Preskill, 2005). The teachers encouraged students' participation and explained that they would be learning a lot by just talking if "you put on your scientist hat," (participate in the SEPs as scientists do). Discussion times permitted students to use the different practices to explore phenomena and use crosscutting concepts to make connections and solve problems they encountered. Due to the lengthy nature of these classroom talks, the teachers usually assigned unfinished work for homework and assisted students as needed through Google classroom.

The challenge of practicing 3D for the teachers at the beginning, as mentioned earlier, was staying consistent in weaving the dimensions and lack of student participation, but it became less of a struggle with continuous implementation. Student participation is essential because it opens an avenue for all viewpoints and allows students to explore ideas and issues in depth from a variety of lenses. Students' participation in these classrooms might have been impacted by large classroom sizes, teacher and student personality, and perception of peers (Abdullah, Baker, & Mahbob, 2012). Additionally, most students were not familiar with this type of learning where "yes, one can make a mistake, but it is alright," (Chelsey). Some of them feared the embarrassment of sharing the wrong answer, and others did not want to risk it and get a bad grade. Luckily, the teachers' constant reassurance that there is no right or wrong answer and that only their reasoning and justification counts, motivated more students to start participating in classroom discussions or in written explanations. Another obstacle was that not all teachers were on board with the 3D teaching, which added to the struggle with student's involvement.

Lack of collaboration among teachers was also seen as an impediment. Literature reveals that American teachers spend more time in the classroom and far less time planning, less than 2% vs. 35% of planning time in high achieving schools (Rosenberg, Daigneau, & Galvez, 2018). Collaborative planning for the "standards-based" or 3D lessons was a resource that the teachers appreciated. Planning together, the teachers could discuss 3D strategies and design lessons together, inform and critique each other, and assess students' thinking and assimilate learning. Co-planning can provide opportunities for mentoring with experience and/or competent teachers, encourage teachers to experiment with new instructional approaches, and co-construct understanding of practice (Johnston & Tsai, 2018). However, the teachers in this study did not collaborate often, and not all their colleagues were willing to or could make the time to meet and plan. Collaboration might have shaped their practices, motivated continuity and growth by shifting the reform focus from individual teachers to the department and to the school. Another reason for the lack of collaboration advanced by the teachers was administrative demands and personality issues.

The teachers implemented 3D starting with deconstructing the standard (created their lesson plans using the core ideas in the standard). They also selected phenomena to reflect the standard based on students' experience and interest. The teachers adapted old activities for students to interact with the dimensions. The students could explore the activities freely and lead discussions, but the teacher controlled the flow of information. The teachers exercised some flexibility.

The recommendation for three-dimensional teaching is for teachers to embrace a metacognitive stance in their teaching, a cycle of reflection and adjustment in what they understand and can do is necessary. Effective 3D teaching and learning will require synergy between the various type of knowledge, knowledge of the dimensions, understanding of the nature of science, and how students learn (Crawford & Capps, 2016). Mere understanding performance expectations of the standard will not provide students with the opportunities to discuss phenomena and propose solutions to problems. Teachers should also train students to start thinking differently about their role in learning and how to use the SEPs a CCCs to help them figure out the core ideas. These practices must be trained and refined continuously. According to Crawford and Capps (2016), even teachers who receive professional development on 3D teaching struggled involving students in science practices.

Question 4.

How do High School Science Teachers Practice FA of 3D Learning?

Science education reform encourages a shift in focus from assessing science ideas to assessing how students figure out phenomena and construct solutions to problems (NRC, 2012). The Framework's vision is for students to acquire knowledge and skills in a sequence of stages that develop their understanding of aspects of three-dimensional practices in each standard. Like the teachers in Herman et al. (2010) study, Chelsey, Paul, and Andria established and communicated their learning goals with students and constructed eliciting questions from the start alongside the lesson activity with these goals. They started from the core ideas and constructed questions or followed a backward design model (Wiggins & McTighe, 1998), to include all the elements of the three dimensions addressed in the standard and to collect evidence of students' learning. In creating their assessments, the teachers deconstructed from rather than build-up to the goal of the lesson. So, the teachers could track where the students were relative to the standard, and the students could learn new concepts by building onto the previous one for incremental learning.

The formative notion of looking forward and continuously assessing where students are relative to the standard or learning goals is possible, with the teacher first mapping the sequence of what students will learn (Herman, 2013). Learning progression can be used to scaffold teachers' design and practice of formative assessment (Furtak, Circi, & Heridia, 2018). Chelsey, Paul, and Andria created their lessons one unit at a time and created assessment tasks concurrently to provide evidence for applicable performance expectations. Nonetheless, it was not clear whether they considered students' pathway to mastery, to guide their formative assessment tasks. Research confirms that teachers struggled to utilize learning progressions to set learning goals, to interpret students' ideas against these goals, and to adjust instruction (Covitt, Gunckel, Caplan, & Syswerda).

Assessment tasks for 3D learning should be composed of multiple components to reflect the interconnectedness of the dimensions, reflect continuous learning by soliciting information along a continuum of results in each grade, be interpretive to evaluate a range of student responses and to guide instructional next step for teachers (NRC, 2014). Chelsey, Paul, and Andria created assessment tasks with a purpose, to mirror the core idea of the standard and to engage students' interest and integrate their reasoning with the concepts. Attention is drawn to the fact that the teachers developed both their assessment tasks and their classroom activities from the core ideas of the standard, as such, the guiding prompts or questions for eliciting ideas and probing critical thinking in the assessment tasks looked like mini versions of classroom activities (NASEM, 2017). However, the assessments tasks were more on the 2D quarters than in the 3Ds. Research suggest that 50% of assessment tasks should be 3D, although they are more difficult to construct at the beginning, it would signal students that the other dimensions are equally important (Underwood, Posey, Herrington, Carmel, & Cooper, 2018). The assessment tasks are rigorous to the extent that they provide students to apple their knowledge to a new situation.

Chelsey, Paul, and Andria employed open dialogue to engage students in discussions and solicited chorus responses to elicit multiple ideas or concepts (Furtak & Ruiz-Primo, 2008), to identify misconceptions and assess students' developing understanding of concepts. The literature supports the use of classroom discussions as an assessment to help the students and teachers adjust (NASEM, 2017). Skill teachers, according to Ford-Connors et al. (2016), "use dialogic exchange with students to both monitor understanding and initiate instructional moves to engage students in deeper explorations of content" (p. 51). These teachers embedded daily assessment conversations or instructional dialogues (Ford-Connors, Robertson, & Paratore, 2016) into an activity currently taking place in the classroom, to help them gather current information about students' conception, language used or communication skills, mental models, and use it to guide and refine instruction (Ruiz-Primo & Furtak, 2006). The conversations were mostly centered on the core ideas and some SEP, but occasionally the teachers used a CCC to help the students explain their thinking. The pattern that emerged (table 11 above) is understandable give that students must master the core ideas in 2D before attempting 3D (Underwood et al., 2018). Student-student-teacher interactions during daily classroom talk improved students' understanding of the dimensions. These discussions were necessary for assisting the teacher to continuously gain insight into students' level of understanding (Bell & Cowie, 2001; Furtak & Ruiz-Primo, 2008).

A theoretical model of learning and its regulation supports the use of descriptive interactive feedback as core to FA. Feedback provides a clear vision of targeted skills, appraises student's current progress, and explains how they could improve (Rushton, 2005; Shepard, 2005). The teachers provided feedback to the students and encouraged them to provide feedback to their peers and to use the evidence to adjust their thinking and their explanations. Good feedback, according to Heritage (2011), should be both actionable and systematically planned to yield targeted information with respect to progress indicators. Chelsey, Paul, and Andria used evidence collected as feedback from classroom talk and activities to take the next step of instruction.

The teachers adjusted their instruction and student learning, including remediation (error analysis and tutorial), re-teaching (revisit concepts as questioning and in warmups), and selecting resources for future instruction. Research indicates that teachers struggle with the next step of instruction (Buck et al., 2010; Herman & Choie, 2008; Herman et al., 2010). Chelsey, Paul, and Andria's primary focus were preparing their students to pass the end of course test whose format is 3D. Therefore, any adjustment in students' learning and their teaching was geared towards this goal, and that may have lessened the struggle. The teachers did not always have time to implement all adjustments intended, though their next step was deliberate (teaching to the test) and driven by the standard and limited in scope.

Chelsey, Paul, and Andria were less enthusiastic about tracking students' understanding of 3D learning because some students were not willing to participate, therefore concealing from the teachers what they think and can do. Active learning occurs when students and teachers actively participate in the learning activity. Abdullah et al. (2012) suggested three factors that might influence student participation, including teacher and student personality, classroom size, and perception of peers. Each of the teachers had more than 30 students of diverse cultural backgrounds in their classrooms. Lack of participation in these teachers' classrooms may had been varying cultural, economic, or linguistic background of students along with the varying degree of comfort levels they bring into the norms of science classroom talk (Abdullah et al., 2012). Students' participation in theses classrooms may had also been influenced by peer perception and students not wanting to be embarrassed by giving the wrong answer or saying something "dumb" in class.

Tracking students' developing understanding of integrated science knowledge can occur informally and formally. The teachers used informal assessment during discussions and questioning from phenomena to formatively assess students' understanding of 3D learning. However, they did not often implement the formal FA because they ran out of time as the discussions lasted longer than expected, and the teachers had to assign them as homework or move them into future lessons.

Another challenge these teachers faced assessing 3D learning was staying consistent in weaving the three dimensions into the performance expectation of their assessment tasks or difficulty creating standard-based formative assessments. One of the teachers acknowledged that because the students were accustomed to answering selective response questions in quizzes and test (which are two-dimensional), she was conflicted to create constructive response tasks that assess three-dimensional learning when most of her students did not attempt or put down IDK (I don't know) or something unrelated. Chelsey, Paul, and Andria wished there were sample questions available to guide them with implementing FA of 3D. Nevertheless, research indicates that there are no verifiable formative assessment strategies (Fick, 2017; Herman, 2013; Underwood et al., 2018) for improving teachers' three-dimensional teaching and learning. This might have been good for the teachers because it compelled them to brainstorm ways to construct their questions and answers from the core ideas rather than being purveyors of questions and answers of others like in the past. An added advantage is that the teachers will become skilled at creating their 3D assessment tasks, given that the process gets easier with practice (Underwood et al., 2018).

The teachers practiced FA starting with constructing questions and tasks from the standard one unit at a time. They collected evidence of students' developing understanding using mostly informal discussions (dialogue and explanations) and occasional formal tasks. Their next step involved using evidence from feedback to adjust their instruction and students' learning. They encountered challenges, including lack of time, unmotivated students, and lack of support and resources. The recommendation for formative assessment of three-dimensional learning is for teachers to move away from assessing only the core ideas addressed in the standard, to assessing the different SEPs students use to obtain information about the core ideas and the CCCs they use to evaluate and communicate their understanding of the core ideas. In this way, the teacher and students will weave the dimensions in their teaching and learning, respectively. The teachers should understand the performance expectations of the lesson and develop their tasks and questions from them. Their assessments should be multicomponent to reflect the interconnectedness of the dimensions. They should use classroom dialog guided by targeted questions to continuously assess students' developing understanding of core ideas and can thus use the feedback to adjust their instruction with respect to students' changing needs. Teachers can assess general and content facts but must also incorporate an equal amount of SEPs and CCCs.

Embracing Three-Dimensional Teaching and Learning

The teachers in this study viewed three-dimensional teaching, not as a laundry list of strategies, instead as a position taken towards teaching science every day (Krajcik, 2015). The findings of this study indicated that the teachers were moving towards embracing the reform message and were making minor adjustments to improve their implementation of 3D teaching.

The answers to research questions two and three, as described above, suggested that the teachers believed that it was essential to integrate the three dimensions in their assessment tasks and classroom activities. Previous studies confirmed that teachers become comfortable in using their assessment tasks, interpreting the outcome, and aligning it with the core ideas when they create it themselves (Yin et al., 2008). The same can apply to Chelsey, Paul, and Andria since they created their assessment tasks. Teachers must shift their thinking from 2D to 3D teaching and stick with the new experience before they could delineate the steps, they must walk for 3D teaching and learning. Acquiring the knowledge and skills for 3D instruction and assessment is tough; it will occur gradually and cumulatively. When teachers acquire new knowledge and skills, they must nurture it through practice and reflection for it to become part of their pedagogical repertoire, to avoid losing it to old and less rigorous ones (Crawford & Capps, 2016).

Chelsey, Paul, and Andria's teaching displayed some evidence of 3D teaching proposed by national documents (NASEM, 2017) including 1) integrating SEP, DCI, and CCC; 2) maintaining flexibility (allowing students to explore and work towards goals set at the beginning of class); 3) working cumulatively in providing continuous support to help students' progressive understanding at each stage of learning; 4) engaging students in daily investigation of phenomena; 5) repeated engagement with important ideas, guidance, and providing opportunities for reflection and 6) providing all students with multiple opportunities to learn science. It is fair to say that although the teachers believed in integrating the three dimensions, implementing it was not easy. They set reminders because they forgot to weave the dimensions or did what they were used to, which is focused on the core ideas. Chelsey, Paul, and Andria created a culture in the classroom of figuring out phenomena and designing and investigating problems for three-dimensional learning. According to Krajcik (2015), “students can build deeper understanding as they grapple with making sense of phenomena or finding solutions to problems. As a result, learners can figure out more complex phenomena or design solutions to more perplexing problems” (P. 16).

Informing Three-dimensional Teaching and Learning for Clarity and Reinforcement.

Formative assessment plays a vital role in guiding instruction and fostering learning (Black & Wiliam, 1998; Herman, 2013). The teachers in this study had a general understanding of the concept of formative assessment and its essential role in 3D teaching and learning. High-stake summative assessments impact how teachers teach and assess learning (Darling-Hammond & Bransford, 2005; Furtak & Ruiz-Primo, 2008). Chelsey, Paul, and Andria used the school district’s Benchmark exams to structure their formative questions to elicit students developing understanding of the dimensions. They were teaching from the test and to the test. The irony, in this case, is that the phrase ‘teaching to the test’ bears a positive connotation than usual. They started with the end in mind and deconstructed classroom activities and formative assessment tasks from the previous year’s exam and the core ideas of the standard, to “collect information of their students learning as it develops” (NASEM, 2017, p. 32). Comparable to Herman et al. (2010), these teachers established and communicated their learning goals and implemented FA to support students’ learning goals for 3D teaching.

Diagnostic evidence informs the next step in teaching and learning (Herman, 2013). However, this adjustment is possible only if the teachers used assessment strategies guided by the core ideas and make valid inferences from the evidence collected. To make the correct inferences and adjustment about 3D teaching and learning, teachers must understand how “students’ understanding and ability to apply scientific concepts and related practices develop and grow more sophisticated over time" (Herman, 2013, p. 7) in each standard. Although these teachers planned their lessons with the core ideas from the start and continuously tracked students learning with graduated tasks, it was not clear whether they understood how to map students' ideas at each stage of students' understanding or they were regurgitating the curriculum standards. The teachers also used rubrics to help them interpret students' work, outlined the expectations of the activity based on the standard, and to thought poignantly about the kind of information needed for the assessment task or activity. The rubrics also helped the students "recognize how far they have progressed and where they still have work to do" (NASEM, 2017, p. 59).

To measure students developing understanding of 3D science learning, assessment tasks must 1) examine how students use SEP in the context of CCC and DCI, 2) use many challenging tasks to provide multiple opportunities to demonstrate learning, 3) elicit diverse and specific information for next step of instruction by the teacher and for students to monitor their progress, and 4) focus on students' conceptual development rather than on right or wrong answer (NASEM, 2017; NRC, 2014). In this study, the first and fourth conditions were evident in all three classrooms. However, condition two and three were limited in the scope of implementation. Difficulty employing challenging tasks and eliciting diverse and specific information was expected, given that the teachers had not been trained and were in the learning stage of practicing 3D (Krajcik, 2015). So, the tasks were not necessarily rigorous, and when they were, the students complained rather than challenge their thinking, or the teacher provided scaffolds. The tasks elicited mainly specific information that addressed the performance expectations of the standards, and the next-step activities elicited related information to help students build their understanding of the standard (Milestone test) rather than pushing them beyond their zone of proximal development. Evidence from other classrooms also revealed that teachers struggled to improve students' integrated science knowledge (Fick & Songer, 2017). Borrowing from Herman's (2013) proposed FA conceptual model, a guide for assembling elements for assessing students' learning of 3D should include, the progression of learning goals, an observable assessment tasks, an interpretive lens, and a feedback loop used to close the gap. The teachers in this study, as mentioned earlier, followed this guideline. Like with Herman et al. (2010), these teachers established and communicated their learning goals and implemented FA to support students' learning goals for 3D teaching.

Shift in Thinking and Teaching

All three teachers believed in creating their classroom activities and assessment tasks from the core ideas of the standard and generating questions to assess students at specific junctures in the lesson to match what they were teaching. In favor of authorship, Yin et al. (2008) explained that teachers are more comfortable using assessment tasks when they are the authors. Contrary to the past, when most teachers were the purveyor of questions and answers of others or textbooks, these teachers were constructors of their questions and answers for the lesson they teach. Despite these teacher's investment in their craft, they also expressed a need for ready-made resources. With the practice of 3D, these teachers were engaged in thinking through the content or subject matter they were entrusted to teach. When teachers are involved in thinking through their lesson, creating and implementing their assessment tasks, it can motivate continuous use, allowing them to interpret and align the concepts to the core ideas (Yin et al., 2008). The result is likely to lead to improved practice and student learning.

Students as key players in three-dimensional teaching and learning

What part do students play or what kind of thinking and learning skills do students need for 3D learning is the question worth pursuing. Some ideas can be borrowed from the literature to focus on students' 3D learning. For example, Kendall et al. (2008) suggested that teachers can involve students in decision-making strategies and metacognitive skills such as "setting their own learning goals, monitoring their progress toward learning goals, and monitoring their thinking processes for accuracy and for clarity" (p.3). Concerning students setting their learning goals of 3D, the teachers in this study identified and shared the learning target with the students and provided them with the context/occasions (phenomena and activities) to work towards each achievement level (NASEM, 2017), however, not every student took advantage of this opportunity. The teachers required students to ask questions to activate their thinking and to answer their questions. For instance, all three teachers responded to students' questions with a reflective toss, sometimes followed with the phrase "what does the standard say." Paul always reminded his students when they were lost (confronted with difficult concepts), to go back to the standard for guidance, "that is your home button" (comparing the standards to the GPS or their phones used for driving).

Students are just as challenged as teachers when acclimating to this new way of teaching. Most students, as indicated earlier, are still believers of the teacher as the possessor of knowledge and should tell them what to do rather than think and exhibit behaviors as scientists. Therefore, they do not trust the system to take the risk of giving the wrong answers or checking their work for correctness, which will allow them to own their learning and set goals toward achieving them.

The teachers believed that the students should be the ones thinking, so they encouraged them to ask questions of what puzzled them and to try and answer it themselves. Students should be resources for each other. They collaborated in their group activities to practice critical thinking of asking questions, making claims, and solving problems when exploring phenomena and culminating activities. Working in groups allowed the students to be creative, think abstractly, or out of the box to figure out the real-world phenomena. Group work helped students communicate what they learned during classroom discussions and presentations, write investigative reports, provide feedback to peers, and complete assessment tasks or classroom activities. Engaging students in a variety of investigations to figure out the real-world based on their interest and experiences assist them with the transfer of knowledge to another context rather than be trapped in boring scripted cookbook experiments (NRC, 2014). Additionally, students would shift from being receptors of information to constructors of knowledge.

Shortcomings of teachers' conceptions, practice, and FA3DL

Andria, Chelsey, and Paul approached the concept of 3D with the belief that it was too complicated, and this may have limited the scope of how far they were willing to explore the concept. Their conception of 3D was that it is a way of teaching by integrating the three dimensions and starting with a phenomenon. However, they fell short in explaining how weaving the three dimensions worked or why it was necessary. Without this knowledge, the teachers missed the opportunity to see how this process could help students to progress towards understanding the dimensions. Also, how the teachers constructed their knowledge and beliefs about 3D influenced their classroom structure, the way they unpacked the standard, and how they practiced 3D (Jones & Leagon, 2014). Although willing, since the teachers saw 3D as tedious and had not experienced the value of this reform, they had to force themselves away from their traditional thinking about instruction to gradually embrace what 3D is and what it can do.

The teachers believed that deconstructing the curriculum standard to plan their lesson and assessment and finding a phenomenon that engages and gets students talking is practicing 3D. Their instruction focused more on getting students to explain and justify their reasoning to teachers' questions (what the standard dictates) and less towards students asking questions, constructing a solution to problems, making connections, and figuring how phenomena work. The teachers were not too concern about how the dimensions were used; they were thinking about it but were not really following through all the times. One of them confessed that "if I see someone walk in my room, I will mention the standard even if I am doing something at my desk. I will stop everything and make that comment… also mention what occurs in another science class to honor the CCC." This quote indicated that the teachers had not truly embraced the practice of 3D and still followed their usual routine. It is likely they did not grasp the necessity of integration and saw it as outside their lessons. Although the teachers included all the dimensions in their lesson plans and displayed them on their boards, they treated the SEP and CCC as add-ons. From talking with the teachers, Paul displayed posters on the wall to remind him to utilize the other two dimensions, Andria inserted a slide in her PowerPoint presentations to remind her to use them, and Chelsey reflected at the end of the lesson whether she had used them or not. Between the SEP and CCC, the teachers were less concerned about SEP because they believed they "had been doing it all along." So, the CCC was considered the most challenging to weave into their lesson and least used. To solve this problem, Talanquer (2019), suggested that teachers might find CCC easy to use if they conceptualized it not as an additional task or content, but as a way for students to “develop the productive ways of thinking that the different CCCs encapsulate” (p. 16).

All three teachers believed that students should do most of the talking and asking the question. However, they were still in control of the information in the classroom. For example, when Chelsey was completing a Punnett square on the board, a student was excited to practice, and she called out, “can I do it? However, Chelsey responded with, “don’t you come up here? The teachers also put notes and resources in Google classroom to make sure the students had the right information after the lesson. This frontloading, however, was counterproductive because some students accessed the information before class, which took away the element of wonderment that drives exploration. All three teachers taught Biology, which is a course with a Milestone exam, and this might have impacted how they implemented 3D.

The teachers also fell short in communicating how FA is integral to 3D teaching, as a piece that improves students’ understanding of CCC and DCI, and their aptitude with SEP (NASEM, 2017). The teachers approached FA3DL as an evaluation of what students knew and could do rather than a collaboration between teacher and student to work towards accomplishing the goal of the lesson (Heritage, 2011). Nevertheless, because the teachers focused mainly on students’ understanding of the DCIs, how students improved their understanding of CCC and facility with SEP were rarely assessed. The rationale, as stated by Paul, was that “the content itself is going to drive the questions and everything that is on the test,” and Andria said, “the core idea determines what we do.” The questions the teachers used were mostly content based, except for one or two questions relating to SEP and CCC. Chelsey articulated more understanding of FA3DL, but it was not evident in her classroom. Teachers will need to gain more knowledge and practice about 3D teaching to facilitate the process of weaving the dimensions to be able to assess all three formatively. The evidence suggests that the teachers were not fully committed to 3D teaching, and it will take support, resources, and practice to turn the tide entirely towards this reform message. A complete shift will depend on another dimension, the teacher, who is willing to change their views of pedagogy, curriculum, and assessment they are responsible for implementing (Rushton, 2005).

Emerging themes for three-dimensional teaching

In exploring teachers' practice of 3D teaching, some common themes emerged. The teachers were leaning towards the belief that 3D teaching was the right way of teaching. Collaboration with colleagues, though seldom occurred, provided resources and motivation to continue practicing. Students' participation allowed sharing their ideas during classroom discussions to move them collectively towards the learning goal. Using phenomena relatable to students' interest and experience, pushed discussions in the classroom forward. Teachers expressed a desire for available resources and strategies for teaching and assessing 3D (although they acknowledged the importance of creating their questions that aligned with the core ideas). There was consensus on time limitations for practicing 3D teaching. In the district where this study was conducted, the teachers experimenting with 3D teaching have pushed for a more extended class period, and the district is considering moving from 50 minutes 7 classes a day to 90 minutes four by four block schedule to give teachers enough time. The teachers exhibited flexibility with time, content, and student ideas and participation, but controlled information in the classroom.

Factors that Influenced How Teachers Formatively Assess Three-Dimensional Learning

All three teachers indicated that planning assessment tasks along with classroom activities based on core ideas were paramount to their formative assessment of 3D, and their efforts were supported by collaborative planning when it occurred. Integrating the three dimensions into the assessment tasks enabled the teachers to assess students' 3D learning. All three teachers stressed the necessity for student participation as a catalyst to guide students' learning and the next step for instruction. No assessment is effective without a clear goal for the lesson. The teachers' understanding of the core ideas and what students should know and could do guided their assessment of 3D. Explanations provided a window into students' thinking. A noticeable shift in how teachers formatively assess students was the absence of quick response devices for selective response questions with an 'end answer,' rather, teachers required students to provide explanations or reasoning for their claims or answer choices to move science talk forward.

Limitations

The instrument in a qualitative study is the researcher (Yin, 2009). Despite debriefing with other researchers, checking with participants, and using multiple sources of data to reduce bias in this study, it was insufficient to maintain neutrality. As a teacher who also works in the county where this study was conducted, sharing a collective experience with the teachers made it challenging to conduct a biased-free research. A constrain with member checking included long time to reply to request and lack of teachers' feedback. Therefore, the results were based on my interpretations to make the judgment call. Conducting a similar study with more than one researcher and in a different context from that of the researcher may yield a more objective result.

Another restriction of this study was the challenge of analyzing whether what teachers were saying and doing was just reproducing what was in the standard, or whether they had a grasp of the process of 3D. It was also not clear whether the findings of this study will be short-lived, and whether the teachers will return to their old ways (lecturing and front-loading) or whether the findings will be long-lasting (given that all three teachers expressed that they were experimenting with this new way of teaching and assessing). Some of the classroom observations of the teachers' planned lessons were interrupted or altered (based on other school activities, teacher choice to switch things around, and absences). So, some of the lessons were not observed in their usual context. The constrain of formative assessment of three-dimensional learning according to the teachers was lack of student participation and students' perception of 3D, time to explore concepts deeply, and lack of collaboration among teachers. In addition, the teacher's perception of and resistance to 3D and administrative perception of good teaching were factors worth noting.

Implications for Future Studies

The findings of how teachers practice 3D teaching and formatively assess students integrated science learning have the potential to guide other science teachers, administrators, professional development agents, and school districts venturing in this new wave of teaching. Evidence from this study could provide important insight towards packaging and tailoring the appropriate grain size for strategies and resources that could lead to significant changes in teaching practices and student learning. Some possible areas of research could emerge from this study. For example, providing continuous professional development to teachers and documenting how they practice alongside learning this new way of teaching. Another avenue can be conducting research to identify science teachers' challenges and resources necessary for the implementation of 3D teaching. A new study can be conducted to expand the field beyond high school to include both elementary and middle school science teachers. Another line of inquiry may include researching effective ways to cultivate productive collaboration practices among teachers to facilitate the implementation of 3D teaching and formative assessment of its learning. More insight is possible with repeating this study with a larger sample size to expand on the findings. A model for formative assessment of three-dimensional learning is suggested below.

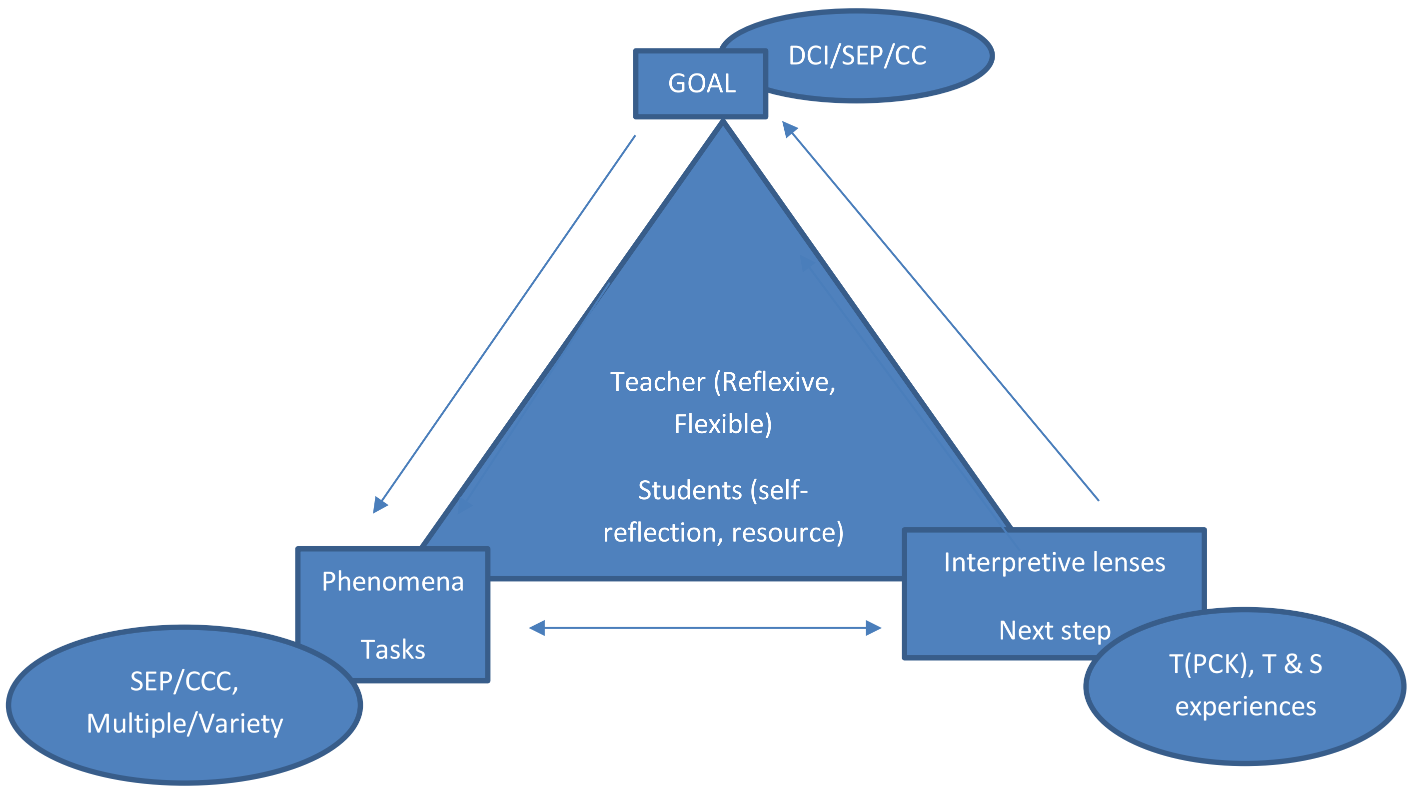

Figure 3. **Formative Assessment of Three-dimensional Learning**

Figure 3 suggests a model for formatively assessing students three-dimensional learning. This model is based on a collaboration between teacher and students, where both parties work from a clear understanding of the success criteria of the lesson, to accomplish the lesson goal (that addresses all three dimensions) and move learning forward. The teacher structures multiple and variety of tasks and phenomena based on this goal for students to explore freely using appropriate SEP and CCC, and they gather evidence or feedback to adjust teaching and learning and to take the next step in learning. This process depends on a critical resource, the teacher's PCK skills and experiences and the students' experiences, as reflexive and reflective lenses.

Conclusion

This study came about because of the new changes occurring in science classrooms and the desire to know how teachers embrace and enact new reforms when the prerequisite knowledge and skills have not yet been fully developed. Enough evidence exists about teachers' lack of understanding and struggles implementing FA. We also know that the concept and practice of three-dimensional teaching are new, with no proposed way to assemble and deliver the message or assess its effectiveness. This study wanted to explore science teachers' understanding of formative assessment and three-dimensional teaching, and the practice of three-dimensional instruction and formative assessment of three-dimensional teaching and learning. Concerning the understanding of concepts, the findings of this study revealed that the teachers had a good understanding of formative assessment but do not yet have a clear understanding of 3D teaching and learning. Regarding the teachers' practice, the findings indicated that they used their experience to implement the three dimensions and were gradually embracing the reformed message. The teachers also pulled from their experiences of formative assessment to gauge integrated science teaching and learning, but with a gradual shift away from evaluating what students know and can do (content), towards a collaboration between student and teacher to accomplish the goal of the lesson (Heritage, 2011) (involving the 3 dimensions).

Few studies are available on science teachers and 3D teaching and learning, and primarily involved an intervention. This study is different in that it is exploratory and involves teachers in the context of their classroom before any specific training on 3D. Insight on teachers' successes and challenges from this study have the potential to guide science teachers and other stakeholders in the journey to understand, implement, and formatively assess 3D teaching and learning.

Teachers will need repeated cycles of training, practice, and reflection to understand the reform message and be able to implement it confidently. In the meantime, teachers will construct their understanding and belief about 3D teaching, which will influence the structure of their classroom, and the way they unpack the curriculum standard and practice it.

REFERENCES:

Abdullah, M. Y., Bakar, N. R. A., & Mahbob, M. H. (2012). Student's Participation in Classroom: What Motivates them to Speak up? *Procedia-Social and Behavioral Sciences, 51*, 516-522

Andersson, C., & Palm, T. (2017a). Characteristics of improved formative assessment practice. *Education Inquiry, 8*, 1-19.

Andersson, C., & Palm, T. (2017b). The impact of formative assessment on student achievement: a study of the effects of changes to classroom practice after a comprehensive professional development program. *Learning and Instruction*, *49*, 92-102.

Anderson, K. T., Zuiker, S. J., Taasoobshirazi, G., & Hickey, D. T. (2007). Classroom discourse as a tool to enhance formative assessment and practice in science. *International Journal of Science Education*, *29*(14), 1721-1744.

Andrews, T. (2012). What is Social Constructionism? *The Grounded Theory Review, 11*(1), 39-46

Aschbacher, P., & Alonzo, A. (2006). Examining the utility of elementary science notebooks for formative assessment purposes. *Educational Assessment*, *11*, 279–203.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science education, 85*(5), 536-553.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy and Practice*, *18*(1), 5-25

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 9–21.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education, 5(*1), 7-74

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, *21*(1), 5-31.

Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In R. W. Tyler (Ed.), Educational evaluation: *New roles, new means (National Society for the Study of Education Yearbook, 68* (2*)*, 26–50).

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher, 33*(8), 3-15

Borko, H., Davinroy, K. H., Bliem, C. L., & Cumbo, K. B. (2000). Exploring and supporting teacher change: Two third-grade teachers' experiences in a mathematics and literacy staff development project. *The Elementary School Journal*, *100*(4), 273-306.

Bransford, J. D. (1979). Human cognition: Learning, understanding, and remembering. *Belmont, CA: Wadsworth.*

Brookfield, S., & Preskill, S. (2005). *Discussion as a way of teaching: tools and techniques for democratic classrooms.* 2nd ed. San Francisco: Jossey-Bass

Brown, J.S., Collins, A. & Duguid, S. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32-42.

Bryman, A. (2008). Of methods and methodology. *Qualitative Research in Organizations and Management: An International Journal*, *3*(2), 159-168.

Buck, G. A., & Trauth-Nare, A. E. (2009). Preparing teachers to make the formative assessment process integral to science teaching and learning. *Journal of Science Teacher Education*, *20*(5), 475-494.

Buck, G. A., Trauth-Nare, A., & Kaftan, J. (2010). Making formative assessment discernable to pre-service teachers of science. *Journal of Research in Science Teaching*, *47*(4), 402-421.

Carless D. (2007). Learning-oriented assessment: Conceptual basis and practical implications. *Innovations in Education and Teaching International*, *44*, 57–66

Chappuis, S. & Chappuis, J. 2008. The best value in formative assessment. *Educational Leadership 65*(4), 14–18

Chappuis, S., & Stiggins, R. J. (2002). Classroom Assessment for Learning. *Educational Leadership*, *60*, 40-43.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Pine Forge Press.

Covitt, B. A., Gunckel, K. L., Caplan, B., & Syswerda, S. (2018). Teachers' use of learning progression-based formative assessment in water instruction. *Applied Measurement in Education, 31*(2), 128-142.

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. London; Thousand Oaks, Calif.: Sage Publications.

Darling-Hammond, L., Adamson, F. (2013). *Developing assessments of deeper learning:* The costs and benefits of using tests that help students learn. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education.

Darling-Hammond, L., & Bransford, J. (2005). *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. Jossey-Bass, An Imprint of Wiley.

Darling-Hammond, L., & Bransford, J. (2007). *Preparing teachers for a changing world: What teachers should learn and be able to do*. John Wiley & Sons.

Darling-Hammond, L., & McCloskey, L. (2008). *Assessment for learning around the world: What would it mean to be internationally competitive? Phi Delta Kappan, 90*, 263–272.

Dewey, J. (1998). *Experience and Education*. New York, NY: Macmillan

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation*, *14*(7), 1-11.

Ebert, E. K., and Crippen, K. J. (2010). Applying a Cognitive-Affective Model of Conceptual Change to Professional Development. *Journal of Science Teacher Education 21*, 371–388.

Feldman A, Capobianco BM (2008) Teacher learning of technology enhanced formative assessment. *Journal of Science Education Technology 17*, 82–99

Fick, S. J. (2017). What does three-dimensional teaching and learning look like?: Examining the potential for crosscutting concepts to support the development of science knowledge. *Science Education*, *102*(1), 5-35.

Fick, S. J., & Songer, N. B. (2017). Characterizing middle grade students' integrated alternative science knowledge about the effects of climate change. *Journal of Education in Science, Environment and Health, 3*(2), 138–138.

Ford-Connors, E., Robertson, D., & Paratore, J. (2016). *Classroom talk as (in)formative assessment*. Voices from the Middle, *23*(3), 50–57.

Furtak, E. M., Circi, R., & Heredia, S. C. (2018). Exploring alignment among learning progressions, teacher-designed formative assessment tasks, and student growth: Results of a four-year study. Applied Measurement in Education, 31, 143–156.

Furtak, E. M., & Heredia, S. C. (2014). Exploring the influence of learning progressions in two teacher communities. *Journal of Research in Science Teaching, 51*(8), 982-1020.

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de León, V., Morrison, D., & Heredia, S. C. (2016). Teachers' formative assessment abilities and their relationship to student learning: findings from a four-year intervention study. *Instructional Science*, *44*(3), 267-291.

Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, *49*(9), 1181-1210.

Furtak, E. M., Pasquale, M., & Aazzerah, R. (2016). How teachers can develop formative assessments that fit a three-dimensional view of science learning. *National Science Foundation*. Retrieved 10/20/17 from http://stemteachingtools.org/link/1800

Furtak, E. M., Ruiz-Primo, M. A., Shemwell, J. T., Ayala, C. C., Brandon, P. R., Shavelson, R. J., & Yin, Y. (2008). On the fidelity of implementing embedded formative assessments and its relation to student learning. *Applied Measurement in Education*, *21*(4), 360-389.

Gearhart, M., Nagashima, S., Pfotenhauer, J., Clark, S., Schwab, C., Vendlinski, T., ... & Bernbaum, D. J. (2006). Developing expertise with classroom assessment in K–12 science; Learning to interpret student work. Interim findings from a 2-year study. *Educational Assessment*, *11*(3-4), 237-263.

Gee, J. P. (2004). Situated language and learning: A critique of traditional schooling. *London: Routledge.*

Glover, C. and Brown, E. (2006). Written feedback for students: too much, too detailed or too incomprehensible to be effective? *Bioscience Education Journal*, *7*(1), 7-3

Greeno, J. G. (2006). Learning as activity. In R. Keith Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 79–86). New York: The Cambridge University Press.

Grooms, J., Enderle, P., & Sampson, V. (2015). Coordinating scientific argumentation and the Next Generation Science Standards through argument driven inquiry. *Science Educator*, *24(*1), 45–50.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, *2*(163-194), 105.

Harris, C. J., Penuel, W. R., D'Angelo, C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., & ... Krajcik, J. S. (2015). Impact of project-based curriculum materials on student learning in science: Results of a randomized controlled trial. *Journal of Research in Science Teaching, 52*(10), 1362-1385

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Heritage, M. (2008). Learning progressions: Supporting instruction and formative assessment. *National Center for Evaluation, Standards, and Students Testing (CRESST). University of California*

Heritage, M. (2010a). Formative assessment: Making it happen in the classroom. *Thousand Oaks*, *CA*: Corwin Press

Heritage, M. (2011). Formative assessment: An enabler of learning. Better: Evidence-based education, Spring 2011. Retrieved from https://www.csai-online.org/sites/default/files/resources/4666/FA_Enabler_of_Learning.pdf

Heritage, M., Kim, J., Vendlinski, T. P., & Herman, J. L. (2008). From Evidence to Action: A Seamless Process in Formative Assessment? CRESST Report 741. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.

Herman, J. (2013). Formative Assessment for Next Generation Science Standards: A Proposed Model. *National Center for Research on Evaluation, Standards and Student Testing (CRESST). Los Angeles, CA: CRESST*.

Herman, J. L., & Choi, K. (2008). Formative Assessment and the Improvement of Middle School Science Learning: The Role of Teacher Accuracy. CRESST Report 740. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.

Herman, J. L., Osmundson, E., Ayala, C., Schneider, S., & Timms, M. (2006). The Nature and Impact of Teachers' Formative Assessment Practices. CSE Technical Report 703. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.

Herman, J. L., Osmundson, E., & Silver, D. (2010). Capturing Quality in Formative Assessment Practice: Measurement Challenges. CRESST Report 770. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.

Hickey, D. T., Taasoobshirazi, G., & Cross, D. (2012). Assessment as learning: Enhancing discourse, understanding, and achievement in innovative science curricula. *Journal of Research in Science Teaching*, *49*(10), 1240-1270.

Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: reconsidering the role of assessment feedback in student learning. *Studies in higher education*, *27*(1), 53-64.

Hock, L. (1961). Classroom Grouping for Effective Learning. Educational Leadership. Retrieved from
http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_196104_hock.pdf

Jasti, C., Lauren, H., Wallon, R. C., & Hug, B. (2016). The Bio Bay Game: Three-Dimensional Learning of Biomagnification. *American Biology Teacher (University of California Press)*, *78*(9), 748-754.

Jenkins, J.O. (2010). A multi-faceted formative assessment approach: better recognizing The learning needs of students. *Assessment & Evaluation in Higher Education 35*(5), 565-576.

Johnston, William R., and Tiffany Tsai. (2018). The Prevalence of Collaboration Among American Teachers: National Findings from the American Teacher Panel, Santa Monica, Calif.: RAND Corporation, RR-2217-BMGF

Jones, M. G., & Leagon, M. (2014). Science Teacher Attitudes and Beliefs. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of Research on Science Education 2*(1), 830-847

Kang, H., Thompson, J., & Windschitl, M. (2014). Creating opportunities for students to show what they know: the role of scaffolding in assessment tasks. *Science Education*, *98*(4), 674-704. Kraus, J. (2012). New Frontiers in Formative Assessment. *Science Teacher*, *79*(4), 74-75.

Kendall, P. C., Hudson, J., Gosch, E., Flannery-Schroeder, E., & Suveg, C. (2008). Child and family therapy for anxiety-disordered youth: results of a randomized clinical trial. Journal of Consulting and Clinical Psychology, 76, 282–297.

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. Educational Measurement: Issues and Practice, 30(4), 28–37.

Krajcik, J. (2015). Three dimensional instruction: Using a new type of learning in the classroom. *Science Scope*, *39*(3), 16-18.

Lauren, H., Lutz, C., Wallon, R. C., & Hug, B. (2016). Integrating the Dimensions of NGSS within a Collaborative Board Game about Honey Bees. *The American Biology Teacher*, *78*(9), 755-763.

Lave, J. (1988). *Cognition in Practice: Mind, mathematics, and culture in everyday life*. Cambridge, UK: Cambridge University Press.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge: Cambridge University Press.

Lee, C., & Wiliam, D. (2005). Studying changes in the practice of two teachers developing assessment for learning. *Teacher Development*, *9*(2), 265-283.

Lemke, J, l. (1990). *Talking Science: Language, Learning, and Values.* Norwood, NJ: Ablex Publishing

Levy, Jack S. (2008) “Case Studies: Types, Designs, and Logics of Inference.” *Conflict Management and Peace Science*, *2*(1), 1-18.

Lortie, D. 1975. *Schoolteacher: A Sociological Study*. London: University of Chicago Press.

Maclellan, E. (2004). How convincing is alternative assessment for use in higher education? *Assessment & Education in Higher Education*, *29*(3), 311-321.

McTighe, J., & O'Connor, K. (2005). Seven practices for effective learning. *Educational Leadership*. *63*(3), 10-17.

Merriam, S. B. (2009). Qualitative research: A guide to design and implementation: Revised and expanded from qualitative research and case study applications in education. *San Franscisco: Jossey-Bass*.

Metin, M. (2013). Teachers' difficulties in preparation and implementation of performance task. *Educational Sciences: Theory & Practice*, *13*(3), 1664-1673

Meusen-Beekman, K. D., Joosten-ten Brinke, D., & Boshuizen, H. P. (2016). Effects of formative assessments to develop self-regulation among sixth grade students: Results from a randomized controlled intervention. *Studies in Educational Evaluation*, *51*, 126-136.

Meyer, C. (1992). What's the difference between authentic and performance assessment? *Educational Leadership*, *49*(8), 39- 40.

Morrissette, J. (2011). Formative Assessment: Revisiting the territory from the point of view of teachers. *McGill Journal of Education/Revue des sciences de l'éducation de McGill*, *46*(2), 247-265.

National Academies of Sciences, Engineering, and Medicine. 2017. *Seeing Students Learn Science: Integrating Assessment and Instruction in the Classroom*. Washington, DC: The National Academies Press

National Research Council. (2001). *Inquiry and the national science education standards*. Washington, DC: National Academy Press.

National Research Council (NRC). 2012a. *Education for life and work: Developing transferable knowledge and skills in the 21st century.* Washington, DC: National Academies Press.

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

National Research Council. (2010). *Preparing Teachers: Building Evidence for Sound Policy*. Committee on the Study of Teacher Preparation Programs in the United States, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

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

Offerdahl, E. G., & Tomanek, D. (2011). Changes in instructors' assessment thinking related to experimentation with new strategies. *Assessment & Evaluation in Higher Education, 36*(7), 781-795.

Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). Thousand Oaks, CA: Sage

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process. *Research and Practice in Technology Enhanced Learning*, *2*(01), 51-74.

Phelan, J., Choi, K., Vendlinski, T., Baker, E., & Herman, J. (2011). Differential improvement in student understanding of mathematical principles following formative assessment intervention. *The Journal of Educational Research*, *104*(5), 330-339.

Phelan, J. C., Choi, K., Niemi, D., Vendlinski, T. P., Baker, E. L., & Herman, J. (2012). The effects of POWERSOURCE assessments on middle-school students' performance. *Assessment Principles, Policy & Practice, 19*(2), 211-230

Putnam, R., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, 29(1), 4-15.

Randel, B., Apthorp, H., Beesley, A. D., Clark, T. F., & Wang, X. (2016). Impacts of professional development in classroom assessment on teacher and student outcomes. *The Journal of Educational Research*, *109*(5), 491-502.

Reiser, B. J., Michaels, S., Moon, J., Bell, T., Dyer, E., Edwards, K. D., & ... Park, A. (2017). Scaling up Three-Dimensional Science Learning through Teacher-Led Study Groups across a State. *Journal of Teacher Education*, *68*(3), 280-298.

Richmond, G. g., Parker, J., & Kaldaras, L. (2016). Supporting Reform-Oriented Secondary Science Teaching Through the Use of a Framework to Analyze Construction of Scientific Explanations. *Journal of Science Teacher Education*, *27*(5), 477-493.

Rivet, A. E.,Weiser, G., Lyu, X., Li, Y., & Rojas-Perilla, D. (2016).What are crosscutting concepts in science? Four metaphorical perspectives. In C. K. Looi, J. L. Polman, U. Cress, & P. Reimann. (Ed). *Transforming Learning, empowering learners: The ICLS, 2,* 970-973

Rosenberg, D., Daigneau, R., & Galvez, M. (2018). *Finding Time for Collaborative Planning-K12- Blueprint.* Education Resource Strategies.

Roth, W.-M. (2005). *Talking science: Language and learning in science*. Lanham, MD: Rowman & Littlefield

Ruiz-Primo, M., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment*, *11*(3-4), 237-263.

Ruiz-Primo, M. A., & 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.

Rushton, A. (2005). Formative assessment: a key to deep learning? *Medical Teacher*, *27*(6), 509-513.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, *18*(2), 119-144.

Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal*, *45*(3), 669-700.

Schwandt, T. A. (2001). *Dictionary of qualitative inquiry (2nd ed.).* Thousand Oaks, CA: SAGE.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., ... & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning. A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, *21*(4), 295-314.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational researcher, 29*(7), 4-14.

Shepard, Lorrie A. 2019. Classroom assessment to support teaching and learning. *The ANNALS of the American Academy of Political and Social Science*. Retrieved from https://doi.org/10.1177%2F0002716219843818

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Stake, R. E. (2000). Case Studies. In N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of qualitative research* (pp. 435-454). Thousand Oaks, CA: Sage.

Stevens, L. R. (2012). *Culturally responsive formative assessment*. Retrieved from Montana State University- Bozeman, Graduate School

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758-765.

Talanquer, V. (2019). Crosscutting Concepts as Productive Ways of Thinking. *Science Teacher*, *87*(2), 16–18. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=a9h&AN=138192430&site=ehost-live&scope=site

Taras, M. (2010). Assessment for learning: assessing the theory and evidence. *Procedia-Social and Behavioral Sciences*, *2*(2), 3015-3022.

Tomanek, D., Talanquer, V., & Novodvorsky, I. (2008). What do science teachers consider when selecting formative assessment tasks? *Journal of Research in Science Teaching*, *45*(10), 1113.

Underwood S., Posey L., Herrington D., Carmel J., and Cooper M. (2018), Adapting Assessment Tasks To Support Three-Dimensional Learning, J. Chem. Educ., 95, 207-217

Vygotsky, S. T. (1978). *Mind in Society: The Development of Higher Psychological Processes.* Cambridge, Mass.: Harvard University Press

Wiggins, G., & McTighe, J. (1998). Understanding by Design. Alexandria, VA: Association for Supervision and Curriculum Development.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment*, (18-40). New York, NY: Routledge.

Wiliam, D. Lee, C. Harrison, C., & Black, P. (2004). Teachers developing assessment for learning. Impact on student achievement. *Achievement in Education*, *11*(1), 49-65

Windschitl M (2002) Framing constructivism in practice as the negotiation of dilemmas: an analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research*, *72*(2), 131–175.

Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education, 45,* 128-136.

Yin, R. (2014). *Case Study Research*: Design and Methods, 5th edition (first edition, 1984), Sage, Los Angeles.

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., ... & Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, *21*(4), 335-359.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education*, *45*(4), 477-501.

APPENDICES

Table 13 Interview Protocol for Formative Assessment and 3D Teaching and Learning

<table>
<tr><td colspan="2">First Interview</td></tr>
<tr><td colspan="2">Greetings, my name is Ms. Koulagna and I am here to talk to you about formative assessment and the new teaching and learning practices we are implementing now across the state. The goal is to capture science teacher understanding and practice of formative assessment in the era of three-dimensional teaching and learning.
Thank you for agreeing to talk with me today. This interview will take roughly 25 minutes, and it will be recorded as auxiliary to what is head. I want to assure you that everything you say today will remain confidential. The recordings and notes will be secured in a safe that is passcode protected and pseudonyms assign to any identifying information. With explanation of confidentiality of this review, I will now turn on the recorder. Do you have any concerns or questions before we begin?</td></tr>
<tr><td>1</td><td>Can you share some background information about yourself including your name, grade level and subject, experience?</td></tr>
<tr><td>2</td><td>How do you make sure you accomplish the goal of the lesson?</td></tr>
<tr><td>3</td><td>How do you determine the kind of task students will be involved in?</td></tr>
<tr><td>4</td><td>How do you determine student progress during the lesson?</td></tr>
<tr><td>5</td><td>How do you provide feedback to students (whole class, groups, or individually)?</td></tr>
<tr><td>6</td><td>How do you respond to student questions?</td></tr>
<tr><td>7</td><td>Do you modify/adjust your instruction as a response to assessing student learning? If yes, how do you do that?</td></tr>
<tr><td>8</td><td>Do you gather evidence of your teaching or student learning? Yes/No why/what</td></tr>
<tr><td>9</td><td>What is your definition of formative assessment? can you share examples/instances of formative assessment in your classroom?</td></tr>
<tr><td>10</td><td>What do you think is the role of formative assessment in teaching and learning?</td></tr>
<tr><td>11.a</td><td>Do you use any resources to support your implementation of formative assessment? if yes, what are these resources and how do you use them?</td></tr>
<tr><td>11.b</td><td>If no, how do you develop formative assessment for your classroom?</td></tr>
<tr><td>12</td><td>Do you have any challenges implementing formative assessment and how do you proceed?</td></tr>
<tr><td>13</td><td>How would you describe three-dimensional teaching?</td></tr>
<tr><td>14</td><td>How do you implement 3D teaching and learning in your classroom?</td></tr>
<tr><td>15</td><td>What challenges or concerns do you have with 3D teaching and learning?</td></tr>
<tr><td>16</td><td>Do you have anything to add about how you formatively assess student evolving understanding?</td></tr>
</table>

Second Interview	
1	What is your definition of FA?
2	How do you practice FA?
3	What are some steps you take after an assessment?
4	Resources for implementing FA?
5	Challenges implementing FA?
6	What is your definition of 3D teaching?
7	How do you practice 3D?
8	How do you formatively assess 3D teaching?
9	What do you consider when structuring your FA task for 3D?
10	What are your challenges of teaching 3D?
11	What are your challenges of assessing 3D?
12	How does teamwork and collaboration help with 3D teaching

Table 14. Synopsis of Empirical studies on formative assessment in science classrooms

Citation	Purpose or Question(s)	Theoretical and Methodological Framework and Assumptions	Methods	Data Analysis	Validity and Reliability	Significant Findings	Limitations	Questions to Investigate Further
Embedded Formative Assessment Tasks								
Tomanek, D., Talanquer, V., & Novodvorsky, I. (2008). What do science teachers consider when selecting FATs?	To investigate science teachers' reasoning associated with task selection of factors used as possible planned FA	-Assessment triangle - Teachers' observational and interpretive skills, and interaction with multiple level of task influence their decision of task selection	-Exploratory case study -24 first and 27 second year pre-service and 41 in-service science teachers -Written assessment probes	Probe response coding at the descriptive and interpretive levels,	Multiple coders for reliability Validity and reliability of instrument is questionable	-Task selection depends on characteristics of task and of student or of curriculum -Training nor experience plays a role	Instrument used as pilot tool.	-What are teachers reasoning about FA? What kind of tasks would have a positive impact on student achievement?
Wiliam*, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement.	To support teachers in planning and exploring their formative assessment practices	Interpretivist views There is no prescribed or unique method of effective classroom practice	Empirical local design 24 science and mathematics teachers Observation, interviews, and standardize test	Quantitative analysis of experimental and comparison group measures with effect size	Validity of measure maintain with use of national test and examination	Improved formative assessment produces substantial gains in standardize test Quality of FA relates to teacher's expertise	Standardization of dependent and independent variables	What kind of support will afford teachers with good FA teaching and better students achievement?
Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., ... & Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change.	To explore whether embedded FA improves students' motivation beliefs that can lead to conceptual change, motivation, and achievement	Learning Theory and Conceptual change FA and conceptual change share similar motivational beliefs	Exploratory quantitative 12 middle school science teachers Videotape of lessons, pre, post-test,	Descriptive statistic with motivation questionnaire HLM with achievement assessment	Exploratory factor analysis efficiency provided reliability and internal consistency between different assessments provided construct validity	Embedded FA in the curriculum had no significant influence on students' motivation, conceptual change, and achievement	Effectiveness of embedded FA depends on teachers' implementation	What is the nature of FA and its feedback potential in today's science classrooms?
Complexity of classroom conversations and response trajectory in formative assessment								
Anderson, K. T., Zuiker, S. J., Taasoobshirazi, G., & Hickey, D. T. (2007). Classroom discourse as a tool to enhance formative assessment and practice in science.	To understand how research efforts to document discursive routines inform participation in scientific inquiry and output on high-stake test	Situated Views of knowing, learning and assessment in context Discourse shapes participation and is itself shaped in the process	-Qualitative Case study -One teacher with 11th and 12th grade students -Videotape of feedback conversation, pre, post- exams, test results, and quizzes from each group	Discourse analysis of feedback conversations, in-class observation and examination of responses to items	Triangulate different discourses and formal gains	Improved teacher facilitation and more productive dialectics with answer rubrics between cycles. Group dynamics influence students' discursive trajectory	Interpretation is limited to researchers' perspective Small sample	What kind of student grouping will improve dialectic discuss among students?
Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts.	To compare the relative utility of the formal and informal functions of four types of assessment prompts in eliciting middle school students' ideas	Framework for FA prompts The utility of assessment prompt as tool for FA tool must elicit multiple levels of student understandings	Multiple methodologies 4 middle school teachers Videotapes of classroom discussions and written responses	Video analysis with each prompt as unit of analysis	Iterative coding the transcript and watching the videotape	conceptions. Prompts elicited multiple conceptions in students' written responses than in discussion	Difficulty comparing students' elicited responses in written and discussion	What are effective ways of using assessment prompt with feedback to elicit students' ideas at a higher level?
Hickey, D.T., Taasoobshirazi, G., & Cross, D. (2012). Assessment as learning: Enhancing discourse, understanding, and achievement in innovative science curricula.	To promote meaningful participation in discursive construction of shared domain knowledge and improve achievement	Situated view of learning Assessment can serve a formative function at one point and a summative at another based on its proximity along a trajectory of curricular activity	Design-based research in a naturalistic setting Middle and High school teachers Observations and interviews (T) Feedback conversation video (S)	Disciplinary discourse and interaction analyzed using sociocultural lens	Used different assessment levels and vary representation of domain knowledge	Enhancement in feedback conversations parallel gains in proximal e and distal measures How formative differ from summative functions is unclear to teachers	Quasi experimental design, instrumentation, and premature termination of project resulting in partial analysis of feedback conversation	What are some strategies that could help redirect science teachers focus more towards FA and less toward summative assessment?
Ruiz-Primo, M., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning.	What is the picture of informal FA in a science inquiry context? Are different levels of informal FA identifiable? Can different levels of	ESRU Framework Informal FA as classroom talk can occur at any level of student teacher interaction	Multiple case studies 4 middle school science teachers	ANOVA with first question, General Linear Model with second and third questions	Multiple raters	Teachers used incomplete cycles 60% of the time. Change in students' post-test results reflect change in	Limited to use of ESRU tool which, does not completely capture the different impact	How will supporting the process of collecting and using assessment results impact teachers' practices of informal FA?

	informal FA practices be associated with levels of student learning?		Videotape of classroom conversations, pre, post-test			teacher's informal FA practices	of the eliciting question	
Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry	To explore the ESRU model to distinguish the quality of teacher's informal assessment practice and to determine whether this quality can be linked to student learning	ESRU framework Assessment conversation allow teachers to elicit and recognize students' conceptions and their communication skills, and use the information to guide their instruction	Multiple case study Three middle school teachers and their students Videotape, FA task, pre, post-test	One-way ANOVA with assessment Linear model rated teachers' average score on graphing	Multiple raters and coders	Quality of teacher's informal FA practice is positively associated with students' learning	Small sample size Multiple context	Can the development of epistemic and conceptual knowledge support teachers' implementation of informal FA?
Perception on and Nature of Feedback								
Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning	To document whether with potential barriers and confusing language of assessment feedback, students would disregard the use of feedback.	Constructivist theory of learning Written feedback must connect with students for effective formative assessment practice	Mixed method 19 higher education students of varying age group, gender, and background Semi-structured interviews of students and questionnaire	Statistical analysis of Likert scale results and interpretation of students' responses	Triangulate methods	Students focus on achieving grades alongside intrinsic motivation.	Limited to quality and quantity of feedback, language used, and impact of subject matter	What are high school science students' understanding of the academic discourse upon which the language of formative feedback is based?
Jenkins, J.O. (2010). A multi-faceted formative assessment approach: better recognizing the learning needs of students.	To explore the use of multifaceted FA in providing students with timely feedback and opportunity to act on it and resubmitting their work	Providing students with immediate feedback of their performance gives them opportunity to correct conception before final submission	Qualitative Action research 32 college students in environmental government course Qualitative Action research	Thematic analysis- data organized inti key themes- identifying students views of feedback and of multifaceted assessment	Thematic analysis	Assessments helped students in managing their study	Study did not measure students' confident in completing assignments with help of scaffold tutorial & guided notes	What is the impact on students' motivation and achievement when given the opportunity to revise their work?
Kang, H., Thompson, J., & Windschitl, M. (2014). Creating opportunities for students to show what they know: the role of scaffolding in assessment tasks.	To document how and why particular forms of scaffolds embedded in assessment tasks guided students' construction of written evidence-based explanations	Framework of instructional scaffolding When teachers engage in effective forms of formative assessment, they provide feedback that addresses students' difficulties	Mixed method 33 first year science teachers Assessment tasks and student written responses	Spearman's correlation coefficient Exploratory Hierarchical multiple regression analysis	Multiple scaffolds used	Combination of scaffolds created opportunities for students to demonstrate scientific understanding through constructing evidence-based explanations.	Non-provided	What combination of scaffolds is needed to improve the quality of high school science students' explanation and learning?
Impact of Professional Development on Teachers' Practices of Formative Assessment Implementation								
Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: a study of the effects of changes to classroom practice after a comprehensive professional development program. *Learning and Instruction*, *49*, 92-102.	To examine changes in twenty-two fourth grade mathematic teachers' formative assessment practices after professional development and its effects on students' achievement	A formative assessment practice whose main goal is to gather information about students' learning implemented by teachers with a mechanistic view may not lead to student achievement	2 teachers Students' pre- and posttest scores	Teachers as unit of analysis ANCOVA	Cronbach's alpha used for internal consistency	Students improved achievement with changes in teachers' formative assessment practices	Teachers' worked in isolation due to random selection with variable context	What is the impact of continuous collaboration on teachers FA practices and on diverse students' achievement
Andersson, C., & Palm, T. (2017). Characteristics of improved formative assessment practice. *Education Inquiry*, 1-19.	explore the characteristics of changes in classroom practice from a combination of formative assessment strategies and the link between the characteristic of these changes and learning opportunities for students.	When students are given the opportunity to adapt and use strategies frequently and gain experience in the process, they improve the quantity and quality of such strategies.	Experimental design 22 - fourth grade teachers and 695 students Interviews observation classroom practices	Descriptive statistic used with each variable Multiple regression analysis	Multiple sources of data	Range of changes in teachers' FA practices from enhancing existing strategy focused on "big idea" to completely changing old practices with new ones	Large number of students and teachers, students' different teachers, duration was very short,	To explore the kind of intervention element (planning or feedback) most successful in improving student learning
Aschbacher, P., & Alonzo, A. (2006). Examining the utility of elementary science notebooks for formative assessment purposes.	To explore the potential of students' science notebook in improving teachers' FA practices and student achievement in elementary science classroom	Constructivist learning theory Teachers' must monitor and diagnose students' understanding of specific concepts to improve such understanding	Mixed method 8 Test and 17 Regular Teachers Student notebook, pre- and post-test, teachers' interviews and observations on nature of guidance provided	Linear regression analysis, Descriptive and thematic analysis	Multiple raters	Notebook as a tool for FA depend on the degree of teacher guidance which itself is dependent on understanding of science content and learning goal	Limited support provided to teachers, the dependent variable	What is the impact of extended professional development that build conceptual understanding of big ideas and learning goal on teacher's use of notebooks?

Buck, G. A., & Trauth-Nare, A. E. (2009). Preparing teachers to make the formative assessment process integral to science teaching and learning.	To understand a teacher's experience with implementation of FA to improve students learning and use this insight to improve teacher educators' practices	Model of classroom FA practice Professional development as a process for educating teachers about FA practice should focus on transforming teaching, learning, and relationships within the classroom	Planning sessions transcripts, lesson plans, interviews(T/S), classroom observations, and student work	Direct interpretation and categorical aggregation	Peer debriefings and member checking help with data validation and interpretation s	Teachers practice of FA evolved, but they first had to forgo their old ways (behaviorist tendencies)	Small sample Students' achievement or experiences with FA cannot be generalized	How does teachers' conceptions of students learning, and sample classroom documents highlight differences in their practices of FA?
Gearhart, M., Nagashima, S., Pfotenhauer, J., Clark, S., Schwab, C., Vendlinski, T., ... & Bernbaum, D. J. (2006). Developing expertise with classroom assessment in K–12 Science: Learning to interpret student work. Interim findings from a 2-year study.	To analyze teachers' evolving expertise with interpretation of student work using portfolio and to identify needed resources and teachers' challenges with weak assessment tasks and criteria	Constructivists views- Framework for assessment expertise Teachers' build assessment expertise by repeated alignment of old and new understanding and practice	Multiple case study 3 science teachers Portfolio and interviews based on initial assessment plans, assessment implementation, interpretation of student work, and final reflections	Hyper RESEARCH - descriptive coding for themes related to interpretation of students' work and those hindering teachers' expertise	Researchers attended all meetings, peer debriefing, primary author read all files for reliability	Teachers slowly embraced improving their interpretation of students' work through integrating new assessment concepts	Weak assessment resources	What is the nature of science teachers' assessment system as it concerns grading and informal assessment?
Lee, C., & Wiliam, D. (2005). Studying changes in the practice of two teachers developing assessment for learning.	To understand the process of teacher change and the development of FA practices that foster this change	Teachers use a combined of quality assessment with quality interpretation and use the evidence to enhance student learning	Qualitative case study 24 teachers Teachers' observations, interviews, videotapes, and document. Researchers field notes	Case summaries- two teachers each as unit of analysis	Detailed investigation of two cases	Slow change in teachers with- Credible evidence, practical ideals, continuous support, reflective intervention, enough time, and flexible implantation	Teachers' nature of change	What can a professional development program do to entice science teachers to embrace FA practices?
Meusen-Beekman, K. D., Joosten-ten Brinke, D., & Boshuizen, H. P. (2016). Effects of formative assessments to develop self-regulation among sixth grade students: Results from a randomized controlled intervention. *Studies in Educational Evaluation*, *51*, 126-136.	To explore the impact of assessment intervention on six grade students' self-regulation, motivation and self-efficacy and whether there is a difference between these different forms of assessments.	Framework for FA	Randomized controlled study 31 teachers and 695 six grade students Self-assessment questionnaires and interviews pre- and posttest assessment	multilevel analysis using Pearson correlation and NVivo	Validate instrument with a pilot study	No impact on peer and self-regulation with use of FA. No significant difference use of peer and self-assessment on students' self-regulation, motivation, and self- efficacy	Multiple context, short duration, and results based on self-report	What is the impact of continuous planning and feedback on students' attitudes towards FA?
Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process	To explore how handheld computers might support improved classroom assessment in science classrooms	Ethnography and grounded theory Teachers' ongoing involvement with design of educational innovation i.e. technology is important support for practice	Qualitative case study 7 teachers Videotape, phone interviews, observations	Descriptive narratives of themes, researchers' synthesis of participant experiences	Multiple interviews and observation	Allow researchers to navigate tension between teachers' views of improving assessment of learning and researchers views of improving assessment for learning	Study was not able to compare the quality of students' learning from co-design innovation to innovation from other methods	How does participation in co-design process influence science teachers' instructional and assessment practices?
Phelan, J., Choi, K., Vendlinski, T., Baker, E., & Herman, J. (2011). Differential improvement in student understanding of mathematical principles following formative assessment intervention. *The Journal of Educational Research*, *104*(5), 330-339.	To address joint challenges of assuring high-quality formative assessments and enabling teachers to use formative assessments more effectively and efficiently.	Effective formative assessment must include not just validated assessment, but also instructional strategies and resources linked to the assessments, as well as professional development	Randomized controlled design 85 teachers & 4091 students Pretest and posttest	Descriptive statistic and two-Level Hierarchical Model unidimensional Rash model factor analysis	Assure many items and variability	Students with better foundation in mathematics profit more with FA. FA is more effective with complex math concepts	Limited time for implementation	How does the use of FA impact students' growth trajectory and variability
Phelan, J. C., Choi, K., Niemi, D., Vendlinski, T. P., Baker, E. L., & Herman, J. (2012). The effects of POWERSOURCE© assessments on middle-school students' math performance. *Assessment in Education: Principles, Policy & Practice*, *19*(2), 211-230.	To determine whether POWER-SOURCE interventions would increase students' performance on bid ideas relative to comparative group	Effective formative assessment must include validated assessments, instructional strategies, and resources tied to these assessments. along with appropriate professional development	Randomized controlled design Field test 19-test and 17 control teachers with their students Extended response and	Descriptive statistic Domain and item analysis Hierarchical linear model	Measures validated with positive alpha	The intervention students out-perform the control group students It had an impact on students' learning	Difficulty recruiting, and high rate of attrition coupled with Poor impact of intervention and students with previous low-test scores	Does POWER-SOURCE students possess a better understanding of basic of key concepts and are they able to transfer the principles to different domains?

			short answer items					
Randel, B., Apthorp, H., Beesley, A. D., Clark, T. F., & Wang, X. (2016).	To estimate the impact of classroom assessment on students' mathematics achievement and involvement, and on teacher's assessment knowledge and skills.	The program should help teachers matching learning targets with assessment methods, providing descriptive feedback, and activating student involvement in learning.	Randomized controlled design 64 teachers and their students Pre- posttest, implementation and students' achievement log teacher assessment practice	Exploratory analysis		Classroom assessment for student learning had no significant impact on students' achievement and teacher practice, but positive for student involvement and teachers' assessment knowledge	Fidelity and variability in implementation Intervention not feasible to teachers'' implementation	How can professional development programs promote teacher ability to engage students in assessment activities that benefit student learning?
Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board Certification.	To explore how teachers' participation in the National Board Certification Program (NBCP) as a professional development learning opportunity can improve everyday classroom practice	Analytic framework Variety of actions goes into teachers' assessment practice and different assessments play different roles in the classroom	Comparative group design 16 middle and high school teachers Videotape of lessons, Ts' written responses about taped lessons, Ts' pre-post interviews and surveys, Ss' surveys & work samples	Survey analysis, T-test	Different raters blindly assigned for generalizability, same rater for familiarity	Test teachers improved their FA practices better than control teacher Chance was brief, shift from grading to formative purposes	Small samples Measures not validated	What is the impact of science teachers FA practices using same professional development model as in NBCPs?
Yan, Z., & Cheng, E. C. K. (2015).	To explore the relationship among primary school teachers' attitudes, intentions, and practices regarding formative assessment.	Theory of Planned Behavior	450 teachers Survey	Rasch scale-descriptive analysis followed by Path analysis		Teacher's intention to practice formative assessment is influenced by instructional attitude, subjective norms, and self-efficacy	Model did not explain teachers reported FA practices well	How will a professional development program arm with sufficient knowledge and skills for FA influence teachers' intensions to practice FA
Buck, G. A., Trauth-Nare, A., & Kaftan, J. (2010). Making formative assessment discernable to pre-service teachers of science.	To explore the preparation of preservice teachers (PST) using a re-conceptualized method course and explore the extent to which this process improves or hinders preservice teachers understanding of FA	Belief in systematic self-examination, reflection, and learning of own practice	Pragmatic action research and field-based case study 30 Preservice Science Teachers (PST) Pre- post-questionnaire, planning meeting transcript, documents, PST interviews, field notes, and field experience	Iterative data collection and analysis Thematic data analysis	Repeated analysis for consensus of interpretations of data	Explicit and re-conceptualized approaches to F and field experience influenced preservice teachers' construction of deeper understanding of FA concepts but they were unable to transfer gains to specific pedagogical strategies	What are some ways teachers could foster or re-establish a trusting relationship with students that will allow them to reveal their alternate conceptions without fear of judgement or punishment?	Relatively short time for field-based study
Feldman, A., & Capobianco, B. M. (2008). Teacher learning of technology enhanced formative assessment.	To examine whether teachers could incorporate technology enhanced FA using Personal Response System (PRS) into their practice and the kind of learning necessary for integration	Constructivist views and Active-learning Pedagogies Action research could improve teachers' practice, their understanding of educational situations, and build new knowledge sought out by others and researchers	Qualitative collaborative action research 8 high school physics teachers in 9th & 12th grade Teacher interviews, collaborative discussions, document- group meeting transcript	Open and thematic coding	Peer debriefing and deep analysis of phenomena	Observe a slow progress from novice to experts It takes a long lime for teachers to change with appropriate support	Limited resources- computer Teachers' lack of expertise in using information technology, the school context, and beliefs about teaching and students	What are the different learning combinations necessary for a teacher to integrate FA into their practice?
Formative Assessment in Relation to Teacher Knowledge								
Herman, J. L., & Choi, K. (2008). Formative Assessment and the Improvement of Middle School Science Learning: The Role of Teacher Accuracy.	To examine the quality of teachers' interpretation of assessment results and the impact of such judgment on student learning	Theory of FA practices Quality FA should include quality tasks, quality interpretation and use the evidence to enhance student learning	Multiple case study 7 middle school teachers and their students Pre-test, post-test, teacher log, students' developmental responses	Descriptive statistics, 3-level Hierarchical Linear Model	ConQuest software provide reliability for pre-post-test measures Triangulate data using multiple vintage points	Teacher's Accuracy to interpret student understanding parallel their ability to improve student learning	Self-selected sample Imperfect measures for rating teachers' interpretation	What are the necessities of assessment accuracy and how can these foster conditions for good practices?
Herman, J. L., Osmundson, E., & Silver, D. (2010). Capturing Quality in Formative Assessment	To study the impact of embedded FA constructs and illuminate differences between assessment	Modern validity theory Teachers' FA strategies elicit	Randomized field study	No analysis instrument was identified	Multiple raters and measures	No relationship between teachers self-reported PCK	Small non-representative sample, psychometric quality and	How does teacher's belief of students' need for substantive feedback influence

Practice: Measurement Challenges.	quality and assessment process.	goal oriented evidence at level of learning required, and provide necessary guide for subsequent instructional decision	39 upper level elementary science teachers Teacher survey, logs, and test, interviews and observation		Coherence and validity of measures	and their assessment practices, quality & process of FA Teachers' assessment capacity is developmental	reliability of measures used for interpretation	provision of next step strategies?
Herman, J. L., Osmundson, E., Ayala, C., Schneider, S., & Timms, M. (2006). The Nature and Impact of Teachers' Formative Assessment Practices.	To explore the quality of teacher assessment practices and its relationship to student learning	Theory of FA practices Quality assessment provides necessary detail for assessing and responding to students' progress with respect to desire goal	Mixed method 9 middle school teachers and their students Observation log of reflection lessons, field notes, pre, post-test, documents	Case summaries of teachers' assessment implementation Hierarchical Linear Model	Observation protocol, Reasoning guide/rubric and Pairing raters for reliability; Multiple sources of data	Teacher possess limited ability to use assessment tools Teachers need time to design and teach new curriculum and new assessments	Lack of teacher collaboration and autonomy	What is teacher experience with implementation of self-designed versus imported FA tasks?
Heritage, M., Kim, J., Vendlinski, T. P., & Herman, J. L. (2008).	To determine the component of variability most likely to score high on teacher knowledge measure and relate findings to teachers in general	Analysis and interpretation is pivotal for effective FA	Generalization study 118 6th grade teachers Assessment tasks	Four-point scoring rubric	Multiple measures	Teachers are better at making inferences about students' level of understanding from assessment information than providing next step for instruction	Instrumentation Imperfect rating	What is the relationship between teachers' ability to formulate next step and adapting their instruction.
Sequencing Learning as Tool for Helping Formative Assessment Practice								
Furtak, E. M. (2012).	To explore students' ideas shared during FA in relation to learning progression and inferences teachers' make of these ideas vis a vis the learning progression	To conduct FA effectively, teachers must possess deep knowledge of content and of common ideas students hold	Qualitative 6 High school teachers Interviews and videotapes	Analytic and descriptive coding	Intercoder reliability from two coders	Teachers could elicit and make inferences on students' ideas based on learning progression Teachers utilized learning progressions to solicit students' misconceptions	Students ideas revealed was dependent on how teacher provided feedback or responded to student ideas	How do teachers' abilities to make inferences about student thinking linked to the presentation of students' ideas in a domain?
Furtak, E. M., & Heredia, S. C. (2014).	To explore how sequencing learning serve as a tool to plan instruction and formative assessment of teachers in two learning communities in two separate studies	Learning is changing participation in communities of practice and is an integral part of the context in which it occurs	Multiple case study 13 teachers from two communities Videotaped of PD meetings, interviews, artifacts	Dedoose,, iterative codes, case summaries	Explicit theoretical assumption, triangulate claims from multiple source of evidence, base findings over multiple years of observation, member checking of claims	Teachers who codeveloped learning sequence used it to plan their instruction and formative assessment while those who were users of the learning sequence had difficult making use of it	The outcome was limited to the context, participants, and communities	Would reversing the role of teachers as creators and user of learning sequence have similar results on coordination of formative assessment?
Furtak et al. (2016)	To explore formative assessment abilities of purposefully selected nine 10th -grade Biology teachers in relation to their students learning	The construct of formative assessment rest on four complementary abilities- designing tasks, asking questions that elicits students' ideas, interpret and provide info and feedback to advance thinking	Product of formative task, videotape of enactment, and pre-and posttest	Analytic coding Hierarchical Linear Model (HLM) and ANCOVA	Two raters for reliability	Teachers design better tasks, accurately interpreted students' idea, and acting on them with learning progression, bur it had no impact on students' achievement	Design of study was unable to separate measures	What is the relationship between teacher's use of learning progression and interpretation of student ideas?
Three-dimensional Learning and formative assessment								
Fick (2017)	To assist student, learn to use the CCC framework to clarify misunderstanding, ask questions of new phenomena and make connection of science ideas across context.	CCC framework supports student's examination of phenomena.	Collaborative study design Researcher and a teacher 70 students Audio & video recordings, pre-post test	Descriptive codes of dimensions and rubrics	Study used a variety of models	CCC frame classroom activities, discussion to examine phenomena, & a conceptual model for understanding	Small sample	Can study apply their understanding of CCC in a context to another?
Reiser et al. (2017)	How does professional development focused on classroom practice help teacher improve proficiency with 3D science?	3D Framework Understanding of core ideas, disciplinary practices, and crosscutting concepts co-develop over time	Exploratory study 24 teacher experts & 22 groups Pre- post survey	Match-pair t-test ANOVA	Confirmatory factor analysis Cronbach's alpha for reliability	Professional development improved sophistication of teacher reasoning about pedagogical scenarios	Difficulty to explore whether and how increase in teacher's expertise leads to changes in classroom interaction and	What type of professional development group interactions is more effective for teacher to grapple with complex scientific practices?

			Constructed response items			involving practices	subsequent student learning	
Richmond, Parker, & Kaldaras (2016)	To examine explanation constructed by teacher candidates as scientific practice for supporting student's 3D learning	Framework of What, How, Why of phenomena	Exploratory study Warm up and lesson plans	Analysis of written discuss	Multiple raters	Teacher candidate's ability to articulate complete and accurate casual explanation for phenomena exist along a continuum	Small sample hinders ability to make correlation between degree of explanation and major or topics	Can the What, How, & Why framework be used to characterize scientific explanations in a different context?
Lauren, Lutz, Wallon, & Hug (2016)	To examine how a collaborative board game about honey bees that simulate worker bees within a colony, could be used to integrate the three dimensions of science education	3D framework Understanding of core ideas, disciplinary practices, and crosscutting concepts co-develop over time	Descriptive study	Visual interpretation of data	Application of understanding of science concept after analysis of data	The game provided a means for students to incorporate scientific evidence & communicated understanding and strategies	Many aspects of scientific practices to coordinate	How would the use of rubric facilitate students understanding of concepts with board games?
Jasti, Lauren, Wallon, and Hug (2016)	How does the environment magnify our exposure to toxicants?"	3D framework Integration of scientific practices	Descriptive study 5 teachers Observation & interviews	Visual interpretation of data	Application of understanding of science concept after analysis of data	Teachers use game mainly to make connections to concepts in the DCI, and to practice modelling and data collection.	Small size Teachers did not implement game as prescribed	What challenges do teachers face implementing a board game on 3D learning?
Harris et al. (2015)	To engage students in science practices of constructing explanations and developing and using models, to demonstrate their understanding of disciplinary core ideas in Earth and Physical science	3D Framework Understanding of core ideas, disciplinary practices, and crosscutting concepts co-develop over time	randomized control trial 40 experimental and 32 control teachers multi- component assessment tasks	Implementation logs Power analysis HLM	Expert review	Treatment group out- perform the control Classroom with low achievers benefit more	Random select makes replication difficult and test teachers had more support	How can teacher fidelity of implementation be measured with accuracy
Fick & Songer (2017)	what alternate integrated science knowledge do eighth grade students demonstrate in response to integrated assessment items	3D Framework Understanding of core ideas, disciplinary practices, and crosscutting concepts co-develop over time	Qualitative case study 8 middle school students 19 assessment questions	Grounded theory based coding Categorical and descriptive analysis	Not identified	Students hold many levels of alternative integrated science knowledge	Small student sample	What are some strategies that can support teacher's development of integrated science knowledge?