

Northwestern College, Iowa

NWCommons

Master's Theses & Capstone Projects

Education

Spring 2023

Enhancing Formative Assessment Through the use of an Embedded Framework

Kyle Ormsbee

Northwestern College - Orange City

Follow this and additional works at: https://nwcommons.nwciowa.edu/education_masters



Part of the [Education Commons](#)

Recommended Citation

Ormsbee, Kyle, "Enhancing Formative Assessment Through the use of an Embedded Framework" (2023). *Master's Theses & Capstone Projects*. 490.

https://nwcommons.nwciowa.edu/education_masters/490

This Article is brought to you for free and open access by the Education at NWCommons. It has been accepted for inclusion in Master's Theses & Capstone Projects by an authorized administrator of NWCommons. For more information, please contact ggrond@nwciowa.edu.

Enhancing Formative Assessment Through the use of an Embedded Framework

Kyle Ormsbee

Capstone Project: An Action Research Project

Northwestern College, Orange City, Iowa

Table of Contents

| | |
|--|----|
| Abstract | 3 |
| Introduction..... | 4 |
| Literature Review..... | 6 |
| Methodology | 15 |
| Research Questions..... | 15 |
| Participants and Research Site | 15 |
| Intervention & Timeline | 15 |
| Variables | 15 |
| Measurement Tool(s)..... | 15 |
| Anticipated Statistical Analysis | 16 |
| IRB | 21 |
| Data Collection | 24 |
| Data Analysis | 24 |
| Discussion | |
| Findings..... | 40 |
| Impact on Teaching and/or Learning | 40 |
| Alignment to Research..... | 40 |
| Limitations | 41 |
| Future Research | 42 |
| Conclusion | 43 |
| References..... | 45 |

Abstract

This action research study was driven by the researcher's interest in formative assessment strategies and methodologies. The purpose of this study was to evaluate the effectiveness of formative assessment on student learning when embedded into practice using a defined framework. Formative assessment has been shown to be an effective teaching strategy that can positively impact student learning. While this is the case, research has also revealed that many educators are implementing formative assessment ineffectively. Educational researchers have proposed embedded formative assessment and the use of formative assessment frameworks as a potential solution to this problem. This study used a mixed-methods approach to assess the impact of embedded formative assessment and the use of William & Thompson's five key strategies framework on student learning. Data was collected in the researcher's middle school classrooms. Data analysis revealed an increase in student learning gains as a result of improved formative assessment practice. The research findings from this study may prove beneficial for educators seeking to improve formative assessment practice. The conclusions provide evidence in support of embedded formative assessment as a methodology to improve formative assessment practice.

Keywords: formative assessment, embedded formative assessment, frameworks

Enhancing Formative Assessment Through the Use of an Embedded Framework

The concept of formative assessment is a prevalent topic of discussion and debate in today's educational world. Proponents view formative assessment as an effective teaching strategy that reforms how teachers utilize assessment in their classrooms in order to increase student engagement and learning. Formative assessment can be defined as a "planned process including a suite of key pedagogical tools used to elicit ongoing evidence of student understanding used by both teachers and students to adjust instruction and procedures within the lesson" (Johnson, Sondergeld, Walton, 2019, p.2409). It provides teachers and students with real-time information that can be used to guide teaching instruction and student learning (Hegazy & Barton, 2017). This type of assessment methodology is in direct contrast to the more traditional summative assessment methodology of simply evaluating learning that has already occurred.

In theory, formative assessment it is an effective teaching strategy that yields greater student learning gains because the educator is using data to inform instruction and provide timely feedback. In support of this belief, there have been several research studies that have supported the use of formative assessment as an effective teaching strategy to increase student learning. While this is the case, there are still practitioners and researchers in education who debate the effectiveness of formative assessment. A primary reason for this belief is the reality that many educators are not effectively implementing formative assessment strategies in their classrooms. In response to this, some educational researchers have developed frameworks to help teachers effectively implement formative assessment into their classroom practice. One such framework is Williams and Thompson's five key strategies framework (Williams & Thompson, 2008). This

framework, which will be discussed in more detail later, provides teachers a working model to help them successfully implement formative assessment into their classroom practice.

The purpose of this study is to evaluate the effectiveness of formative assessment on student learning when implemented within a defined framework. In order to do this, the researcher will embed Williams and Thompson's 5 key strategies framework into their classroom practice (Williams & Thompson, 2008). As a result, the primary research question for this study will be: Does formative assessment have a greater impact on student learning when teachers embed it into classroom practice using William & Thompson's five key strategies framework? By embedding formative assessment into classroom practice using Williams and Thompson's five key strategies framework, student learning will increase at a greater rate than through the informal use of formative assessment. This research will help to show that formative assessment is an effective teaching strategy when it is understood and embedded into classroom practice.

In order to gain an in-depth understating of formative assessment, a review of the literature was conducted prior to testing. This literature review will be organized into four themes. Theme one will include a description of formative assessment. Theme two will discuss the key components and strategies of formative assessment. Theme three will address the debate around formative assessment and why some educators caution against its use. This will include a review of various studies on the topic. Finally, theme four will discuss proposed formative assessment frameworks that can be embedded into classroom practice. This will include an in-depth discussion of Williams and Thompson's five key strategies framework.

Research for this action research study was done utilizing resources from the DeWitt Library at Northwestern College. All included resources are from peer-reviewed journals that were published within the last ten years. These resources included studies regarding the history

and current status of formative assessment, perceptions about formative assessment, various formative assessment strategies, formative assessment professional development, and formative assessment frameworks. These studies were used to form a greater understanding of formative assessment and provide the basis for this action research study.

Review of the Literature

Formative assessment is a concept that can be traced back to the 1960s. During this time, educational scholar Michael Scriven introduced the idea of “formative evaluation” (Scriven, 1967). For Scriven, formative evaluation differed from the commonly practiced summative evaluation in that it was not merely a final judgement about learning but “may have a role in the on-going improvement of the curriculum” (Scriven, 1967, p.41). Quickly after this idea of formative evaluation was proposed, it was further defined and extended by Benjamin Bloom. Bloom stated that formative evaluation should be used to provide feedback at each stage in the teaching-learning process (Bloom, 1969). He proposed that this evaluation could be done using brief tests that could be ungraded and then used as aides in the teaching and learning process (Bloom, 1969). This began a shift in educational practice where assessment was viewed as a tool for learning as opposed to just a tool that measured learning.

Since its inception in the 1960s, the theory of formative assessment has continued to evolve. Today, there are several proposed definitions and interpretations of what constitutes effective formative assessment. One of the predominant definitions of formative assessment used by educators today was proposed by researchers Paul Black and Dylan Williams. These leaders in the field of formative assessment research state that “practice in a classroom is considered formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make instructional decisions that are better founded than

the decisions they would have taken otherwise” (Black and William, 2009). This definition supports the primary purpose of formative assessment in that it is a strategy where educators utilize assessment data to make instructional decisions that better support student learning. In contrast, summative assessment utilizes assessment data to provide a “final decision about students’ learning, for example, for selection, certification, or accountability purposes” (Wools & Veldkamp, 2019, p.3).

Building upon this definition of formative assessment, researchers began working to identify effective formative assessment strategies for classroom practice. In a seminal study on formative assessment strategies, Dylan Williams and Marnie Thompson were able to identify the following five key strategies:

1. Clarifying, sharing, and understanding learning intentions and criteria for success.
2. Engineering effective classroom discussions, activities, and learning tasks that elicit evidence of learning.
3. Providing feedback that moves the learner forward.
4. Activating learners as instructional resources for one another.
5. Activating students as owners of their own learning (Williams & Thompson, 2011).

Together, these strategies provide a framework that educators can use to implement formative assessment into their classroom practice. While this framework provides a foundation for effective formative assessment implementation, it is essential that educators have a clear understanding of these strategies.

Effective formative assessment is an evidence-based process of gathering information on three questions: Where am I going? How am I doing now? Where do I go next? (Beesly et al., 2018). For students to know where they are going, teachers must understand, clarify, and share

the learning intentions and criteria for success. In a formative classroom, what can this look like? To begin with, teachers need to curriculum plan with the end in mind. Williams advocates that teachers use a two-stage process to complete this step (2011). In the first stage, teachers clarify the learning goals themselves by identifying what is worthy and requiring of understanding. In the second stage, teachers identify and establish evidence that will count as success criteria. Once these two steps have been taken, the teacher can now begin the process of selecting activities that will help students meet the learning goals. This process of establishing clear learning goals facilitates the “assessment of the most relevant knowledge and skills, which in turn provided more valid and reliable data about students’ learning needs and thus the possibility to make more valid and reliable decisions about how to act to support all students’ learning towards the learning goals” (Almuntasheri, 2016, p.13).

Once teachers have identified and mapped out the learning goals and criteria for success, they must communicate them well. To do this, teachers must embed the learning goals into their classroom practice (Beesly et al., 2018). Learning goals should be included in each lesson plan, displayed daily, and written in simple language (Williams, 2011). To communicate criteria for success, teachers can use annotated student examples and rubrics. By doing these things, teachers provide students with a clear understanding of what they need to learn and what successful learning looks like. This was validated in a 2017 study titled Characteristics of improved formative assessment practice (Andersson & Palm, 2017). In this study, researchers concluded that “clearer goals facilitated the assessment of the most relevant knowledge and skills, which in turn provided more valid and reliable data about students’ learning needs and thus the possibility to make more valid and reliable decisions about how to act to support all students’ learning towards the learning goals” (Andersson & Palm, 2017, p. 118).

Another key strategy is that teachers engineer effective classroom discussions, activities, and learning tasks that elicit evidence of learning (Thompson, 2011). Once teachers know what is essential for students to learn, teachers should possess a toolbox of varying instructional approaches to elicit evidence of student understanding and to meet their needs (Brink, 2017). According to researchers, one of the most important elements of this strategy is questioning. Numerous studies on teacher discourse have concluded that the primary form of questioning used in the classroom is the initiation-response-evaluation (IRE) approach (Dini et al., 2020). This common type of questioning acts as a channel for “closed” questions and can unintentionally restrict student thinking (Dini et al., 2020). In a formative classroom, teachers need to avoid IRE questioning and seek to ask probing questions that lead to discussion.

In a 2020 study that analyzed the formative assessment strategies of experienced science teachers, the importance of eliciting evidence through effective questioning was emphasized. In this study, researchers categorized question types into the following categories: general, specific, leading, and probing sequences (Dini et al., 2020). General questions were not specific to anything the student said. Specific questions addressed an idea in student answers. Leading questions moved students toward a predetermined answer. Finally, probing sequences involved two or more deliberate questions that were asked in a sequence (Dini et al., 2020). The researchers reported that “teachers’ use of probing sequences of specific questions helped them better understand their students’ thinking, informed instructional decisions, allowed students to clarify their thinking, and provided opportunities for other students to connect their understanding” (Dini et al., 2020, p.293). Teachers in the study who utilized this type of questioning more frequently than others were better able to elicit and interpret student evidence

in order to advance learning (Dini et al., 2020). This highlights the importance of incorporating quality questioning and discussion in a formative classroom.

Williams' third key formative assessment strategy is to provide feedback that moves the learner forward. For this to happen, feedback must be understandable and have a motivational element to it rather than a strictly negative connotation (Williams, 2011). In addition, feedback should be given as quickly as possible, and provide actions that can be used by students to drive their own learning (Hegazy & Barton, 2017). This use of feedback enhances the relationship between teachers and students in the learning process and helps foster an empowering student-centered learning environment (Naghdi-pour, 2017). Unfortunately, research has shown that many educators struggle to utilize feedback in this manner.

In a seminal study on teacher feedback, researchers Avraham Kluger and Angelo DeNisi concluded that most feedback students receive has either no impact or is counterproductive to learning (1996). A primary finding in the study was that feedback focused on the person rather than the quality of work. This type of feedback typically led to comparison with others and had a demoralizing effect (Kluger & DeNisi, 1996). In contrast, feedback that effectively moved the learner forward told them not just what to do to improve but also how to get there (Kluger & DeNisi, 1996). This was also confirmed in a 2018 study that investigated the effects of formative assessment on student achievement. When commenting on the role of feedback in formative assessment, researchers stated that "formative assessment feedback should be designed in a way so that teacher and student receive a clear picture of where the learner is going, where the learner is right now and how to get there" (Pinger et al., 2018, p.63). The results of these studies reveal the importance of providing quality feedback that moves the learner forward. When this is done well and in a timely manner, feedback can be a highly effective formative assessment strategy.

The fourth key formative assessment strategy is to activate learners as instructional resources for one another. When teachers create opportunities for students to monitor and regulate their learning, then learning can significantly increase (Williams, 2011). By doing this, teachers eliminate their tendency to force all interactions to go through them which slows down the learning process. Instead, students are empowered to help each other consistently have conversations about where they are and where they need to go with their learning. To do this, Williams suggests that teachers create routines in which students act as instructional resources for one another in a serious and structured way (Williams, 2011). One highly efficient and effective way to do this is to incorporate think-pair-share activities. By consistently doing these activities, a collaborative culture is created where all students engage in learning conversations and become instructional resources for one another. This is important because multiple sources have confirmed that peer-to-peer assessment can have a significant impact on student learning (Brink & Bartz, 2017).

The final strategy of Williams and Thompson's framework is to activate students as owners of their own learning (2011). The power behind this strategy is that it helps students develop self-regulatory skills, a growth mindset, and confidence. When this is achieved, student learning is likely to increase. To do this, teachers need to help students understand where they are and where they need to go with the learning intentions and criteria for success. Some strategies to help with this could include providing and discussing unit planning overviews, setting out learning checkpoints, and providing exemplars of varying performance levels (Williams 2011). Using teacher-designed rubrics has also been shown to be an effective strategy to help activate students as owners of their own learning (Bernard, et al., 2019). Creating

opportunities for students to perform self-assessments help them learn independently and will likely lead to a more productive learning process over time.

These approaches support the implementation of high-quality formative assessment because they engage teachers, learners, and peers in the process of collecting evidence of student learning to inform the teaching and learning process. When this is the case, multiple sources have confirmed that the use of formative assessment can significantly impact student learning. For example, a 2017 study conducted by the Regional Educational Library on the use of formative assessment supported the belief that formative assessment has a positive effect on student learning. Researchers in this study reviewed 23 rigorous formative assessment studies. After completing their review, researchers concluded that “Overall, formative assessment had a positive effect on student academic achievement. On average across all the studies, students who participated in formative assessment performed better on measures of academic achievement than those who did not” (Klute, Apthorp, Harlacher, Reale, 2017, p.2).

While numerous studies have supported the use of formative assessment as an effective teaching strategy, there are still detractors that suggest formative assessment is not as beneficial as some would like you to believe. One of the primary reasons for this belief is that numerous definitions and ideas of formative assessment exist. This reality makes the implementation and evaluation of formative assessment a difficult task (Wools & Veldkamp, 2019). When conducting a review on the state of formative assessment, researchers Gordon Stobart and Therese Hopfenbeck state “a concern for the review is that current definitions of formative assessment cover a wide range of teaching and learning practices while research design often lack an action theory, often accompanied by a lack of systematic data collection” (2014, p. 6).

These issues led them to the conclusion that the positive effects of formative assessment on learning have been oversold (Stobart & Hopfenbeck, 2014).

To go along with this belief, there are also studies that have pointed to the mixed effects of formative assessment in practice (Schildkamp, 2020). These findings have caused some researchers to suggest that formative assessment is a complex undertaking that is not being used properly in the classroom. “Research has suggested that many attempts to implement formative assessment have produced disappointing results because approaches were not used to their full potential but were rather reduced to mechanistically applying a set of principles” (Schildkamp, 2020, p.3). For these reasons, researchers have stressed the need to provide teachers with clear definitions and ways in which to implement formative assessment well (Andersson & Palm, 2017).

In response to research indicating that many teachers struggle to implement formative assessment effectively, researchers have proposed the need for teachers to embed formative assessment into classroom practice. Embedded formative assessment can be defined as “both an instrument and a process, whereby evidence is purposefully gathered, judged, and used by teachers, students, or their peers for decisions about actions to support student learning” (Wools & Veldkamp, 2019, p.716). In theory, it can lead to the more purposeful use of formative assessment because it is integrated into daily classroom practice and relies on multiple methods of assessment. These methods of assessment can include informal assessments, planned-for-interactions, and curriculum-embedded assessments (Wools & Veldkamp, 2019).

To help embed formative assessment into classroom practice, researchers have suggested the use of frameworks as an instrument for guidance and implementation. One such framework is Williams and Thompson’s five key strategies framework. Researchers have touted the use of

this framework because of its simplicity, use of critical formative assessment strategies, and emphasis on the “big idea” of using evidence to adjust instruction to better meet the identified student learning needs (Andersson & Palm, 2017). Utilizing this framework can help teachers integrate formative assessment into classroom practice to make up a unity (Williams, 2008). When this is achieved formative assessment becomes embedded into classroom practice and is used by teachers, students, and peers to dramatically increase student learning.

While this can be the case, there is still little empirical evidence to support these claims. That is because the practice of embedding formative assessment into classroom practice using frameworks can be a very complex undertaking (Andersson & Palm, 2017). While this may be true, that is where this study finds its purpose. In this study, the researcher will be seeking to evaluate the effectiveness of formative assessment on student learning when implemented within a defined framework. The framework that will be embedded into classroom practice will be Williams and Thompson’s five key strategies framework. As a result, the primary research question will be: What is the impact of formative assessment on student learning when embedding Williams & Thompson’s five key practices framework into classroom practice? The second research question will be: Does formative assessment have a greater impact on student learning when it is embedded into classroom practice as opposed to when it is informally used by teachers? By embedding formative assessment into classroom practice using Williams and Thompson’s five key strategies framework, the researcher believes that student learning will increase at a greater rate than through the informal use of formative assessment. This research will add to the discussion on ways to effectively integrate formative assessment into classroom practice.

Research Context

The primary research question for this study is: Does formative assessment have a greater impact on student learning when teachers embed it into classroom practice using William & Thompson's five key strategies framework? The dependent variable in this study is the measurement of student learning. Student learning will be measured using a dependent samples T-test of pre-test and post-test scores. The independent variable for this study is the embedding of Williams & Thompson's five key strategies framework into classroom practice. This framework was embedded during phase two of the study and will be compared to the informal use of formative assessment strategies during phase one of the study. Some of the controlled variables or constants included the same setting, participants, instructional days allotted to each phase of the study, the use of a pre-test and post-test during each phase, the use of the same application for test taking, etc.

The research setting for this study was conducted in the researcher's 7th-grade science classrooms at Valley Christian Middle School. Valley Christian Middle School is a private school that is part of the greater Valley Christian School District. Valley Christian Schools is a Preschool-12th grade school system that has around 1,000 total students. There are separate elementary, middle, and high school campuses. Tuition at Valley Christian Middle School is \$13,750.00 per school year. The middle school is also a 1:1 school with each student and teacher utilizing an iPad daily. The classroom is located on the middle school campus, and it is where much of the research will be conducted. Some learning activities during the study will briefly take place outside of the classroom but will not involve any additional participants.

The research participants for this study were the researcher's 7th-grade science students. There were 69 total participants from four different class periods. Of the 69 total participants, 35

were female and 34 were male. Of the 69 participants, nine students have a 504 plan and require general classroom accommodations. These accommodations could include things like sitting the students near the source of instruction, providing visual learning aids, allowing more time for test completion, allowing fidgeting and frequent breaks, etc. The learning goals and expectations are not modified for these students. Of the 69 participants, one is also an international student from China who speaks English as a second language. The researcher was given an exemption by the Northwestern College Institutional Review Board to conduct this research with the participants as a part of everyday classroom practice.

Methodology

Each of the 69 participants from the four separate classroom periods was a part of the experimental group in this study. There were no separate control groups as in a traditional experimental study. The purpose of the study was to measure the impact of formative assessment on student learning when the teacher embedded a formative assessment framework into classroom practice. In order to do this, the research was organized into two separate phases. Each phase lasted one school week for a total of 10 days.

Phase One:

In phase one, the researcher taught through a weeklong unit on invertebrates. This unit did not involve the researcher intentionally embedding a formative assessment framework into classroom practice. Formative assessment strategies were informally used by the researcher. During this unit, students took a digital pre-test on the Canvas app at the beginning of the unit to measure current understanding. Results from this pre-test were used by the researcher to help identify areas of strength and weakness. Over the next three class periods, the researcher taught three different invertebrate lessons utilizing various learning activities. Formative assessment

strategies were utilized during these lessons, but they were implemented informally. One such formative assessment strategy was having students complete an invertebrate Socrative. Socrative is a digital formative assessment tool like Kahoot and Quizlet. Student scores remained anonymous, and results were used by the researcher to adjust instruction. At the conclusion of the unit, students were required to complete a digital end-of-unit summative assessment on the Canvas app. This assessment consisted of 16 questions and was worth 11 total points. The results of this assessment and the pre-test were compiled into data tables for each class period to be used for quantitative data analysis.

Phase Two:

During phase two of the study, the researcher taught through a weeklong unit on vertebrates. The timeframe and setup for this phase were identical to phase one. On day one, students were asked to take a digital vertebrate pre-test on the Canvas app. On days 2-4, lesson plans were created and delivered to help students engage with the unit objectives. On day five, students took a digital summative vertebrate quiz on the Canvas app to assess their level of understanding. Results from the pre-test and end-of-unit summative assessment were compiled into data tables to be used for quantitative data analysis. The biggest difference of phase two of the study was that the researcher intentionally embedded William & Thompson's five key strategies framework into classroom practice. This was done with the goal of improving the researcher's use of formative assessment in order to positively impact student learning. Below are descriptions of how the researcher embedded this formative assessment framework into classroom practice during this phase.

Prior to the beginning of the unit, the researcher developed unit a plan that intentionally embedded Williams & Thompson's five key strategies framework into each lesson. Again, this framework includes the following formative assessment strategies:

1. Clarifying, sharing, and understanding learning intentions and criteria for success.
2. Engineering effective classroom discussions, activities, and learning tasks that elicit evidence of learning.
3. Providing feedback that moves the learner forward.
4. Activating learners as instructional resources for one another.
5. Activating students as owners of their own learning (Williams & Thompson, 2011).

In order to embed strategy #1, the researcher planned the unit with the end in mind and determined learning goals and criteria for success. Once learning goals and criteria for success were identified, the researcher designed a pre-test and post-test that accurately assessed students on these objectives. The learning goals and criteria for success were also included in each of the unit lesson plans, and learning tasks were developed with these in mind. Finally, the learning goals and criteria for success were communicated at the beginning of each lesson so that students interacted with them daily.

In order to embed strategy #2, the researcher incorporated a variety of learning tasks that elicited evidence of student understanding into each lesson. The ungraded vertebrate pre-test was taken by students at the beginning of the unit in order to identify areas of strength and weakness. Classroom discussions played a vital role in eliciting evidence of student understanding. During these discussions, probing questions were asked to help take students through ESRU cycles (Menon, 2018). In ESRU cycles, the teacher elicits a question, the student responds, the teacher recognizes the student's response, and the teacher uses the response (Menon, 2018, p.44).

Think-pair-share discussions were facilitated on multiple occasions during each lesson to achieve this. Evidence from these discussions were then used to reteach challenging concepts or misunderstandings. Digital formative assessment tools such as Socrative and Blooket were also utilized to give students a fun opportunity to convey their understanding of the learning goals. Evidence from these digital tools were compiled into data tables with problem questions being identified. Once problem questions were identified, table partners would discuss these questions to help each other achieve understanding.

For key strategy #3, feedback was delivered through a variety of methods. One form of feedback was through classroom discussions. The evidence elicited from these classroom discussions was used by the researcher to provide immediate feedback that moved learners forward. Another form of feedback came through written feedback on assignments. This feedback would be given as quickly as possible and also let students know the steps that could be taken to improve. This was done in response to a study done on feedback where researchers concluded that “feedback should be given as quickly as possible and provide actions that can be used by students to drive their own learning (Hegazy & Barton, 2017). Another form of feedback was given using the digital formative assessment tool Socrative. The vertebrate Socrative was designed so that students would receive immediate feedback after answering each question. After answering each question, students would be shown the correct responses in comparison with their responses. These different forms of feedback provided students with plenty of information to take ownership of their learning and move forward.

Key strategy #4 involves activating learners as instructional resources for one another. In order to do this, peer assessment was utilized as an instructional strategy. This was done because research has shown that students value peer feedback in order to inform their next learning steps

(DeLuca, et al., 2018). During this phase, students were tasked with creating a vertebrate groups mind map diagram. Students were given a set of instructions and a grading rubric to help them complete this diagram. Prior to submission, the researcher designed a peer assessment rubric to be completed by each student. Students were tasked with peer reviewing their table partners' mind map based on the peer assessment rubric. Upon completion, students discussed the peer assessment rubrics together along with ways to possibly improve their mind maps prior to submission. To complete the process, students then quizzed each other based on the information in each other's mind maps to help them prepare for the vertebrate summative assessment. Another form of embedding this strategy came through the variety of think-pair-share discussions used during this phase of the study.

For the final key strategy, the researcher sought to activate students as owners of their own learning. A strategy that the researcher used to achieve this was to have students write down the unit learning goals, criteria for success, and unit overview. By doing this, students would know exactly what the expectations were and what they needed to do to get there. Students then wrote down goals for the unit and steps that could be taken to achieve those goals. Another strategy that was used by the researcher was providing a rubric for the vertebrate mind map. This rubric was explained to the students, and they were given opportunities to self-assess their work based on the rubric prior to submission. By having students take these steps the researcher embedded key strategy #5 into classroom practice.

At the conclusion of the unit, students were required to complete a digital end-of-unit summative assessment on the Canvas app. This vertebrate quiz consisted of 19 questions and was worth 14 total points. The results of the vertebrates quiz and the pre-test were compiled into

data tables for each class period. This quantitative data will be analyzed in comparison with the invertebrates assessment data.

Data Collection

For data collection, the researcher selected an embedded design mixed-methods approach (Efron & Ravid, 2020). This was done to allow the quantitative data to have priority over the qualitative data. During phase one of the study, the quantitative data included the results of a pre-test, Socrative session, and post-test. The phase one pre-test, titled the invertebrates pre-test, assessed students' current understating of invertebrate characteristics and phyla. This digital pre-test consisted of 14 total questions and included a mix of multiple choice, true or false, and multiple answers questions. Upon completion, participant scores were stored in the Canvas app. Along with storing each participant's score, this app displays the class average, high and low scores, and the standard deviation. This data was then transferred from Canvas into Microsoft Word data tables by the researcher.

For Socrative, the researcher designed a 16-question quiz on invertebrate characteristics and phyla. These 16 questions included a mix of multiple choice, true or false, and multiple answers questions. On day #3 of phase one, each participant accessed the Socrative app on their iPads and took the quiz at their own pace. Student names and scores remained anonymous to avoid participant comparisons. Once each participant finished the Socrative quiz, the researcher generated a classroom report created by the Socrative app. This report included individual student scores, class averages, and the percentage of correct responses to each question. This data was then transferred to Microsoft Word data tables by the researcher.

The final data collection method during phase one was the invertebrates post-test. This digital post-test consisted of 16 total questions but was scored out of 11 total points. Some

questions were only worth half a point which accounted for this difference. The post-test included a mix of multiple choice, true or false, multiple answers, and short answer questions. Upon completion, participant scores were stored in the Canvas app. Along with storing each participant's score, this app displays the class average, high and low scores, and the standard deviation. This data was then transferred from Canvas into Microsoft Word data tables by the researcher.

During phase two of the study, both quantitative and qualitative data were collected. The phase two quantitative data collection methods were identical to those used during phase one. This included the data from a pre-test, Socrative session, and post-test. The phase two pre-test, titled the vertebrates pre-test, assessed students' current understating of vertebrate characteristics and classes. This digital pre-test consisted of 14 total questions and included a mix of multiple choice, true or false, and multiple answers questions. Upon completion, participant scores were stored in the Canvas app. Along with storing each participant's score, this app displays the class average, high and low scores, and the standard deviation. This data was then transferred from Canvas into Microsoft Word data tables by the researcher.

For Socrative, the researcher designed a 14-question quiz on vertebrate characteristics and classes. These 14 questions included a mix of multiple choice, true or false, and multiple answers questions. On day #3 of phase two, each participant accessed the Socrative app on their iPads and took the quiz at their own pace. Student names and scores remained anonymous to avoid participant comparisons. Once each participant finished the Socrative quiz, the researcher generated a classroom report created by the Socrative app. This report included individual student scores, class averages, and the percentage of correct responses to each question. This data was then transferred to Microsoft Word data tables by the researcher.

The digital vertebrates post-test consisted of 19 total questions but was scored out of 14 total points. Some questions were only worth half a point which accounted for this difference. The post-test included a mix of multiple choice, true or false, multiple answers, and short answer questions. Upon completion, participant scores were stored in the Canvas app. Along with storing each participant's score, this app displays the class average, high and low scores, and the standard deviation. This data was then transferred from Canvas into Microsoft Word data tables by the researcher.

At the end of phase two, qualitative data was collected to create a more well-rounded data picture. This qualitative data was collected using a student survey. The survey was created using Google Forms and utilized a five-point Likert scale. The survey consisted of the following five questions:

1. The use of formative assessment prepared you well for the invertebrate and vertebrate quizzes.
2. The teacher used formative assessment strategies more consistently and effectively during the vertebrate unit in comparison to the invertebrate unit.
3. I felt more confident and prepared when studying for the vertebrate quiz than I did for the invertebrate quiz.
4. I felt more confident while taking the vertebrates quiz than I did while taking the invertebrate quiz.
5. I had a better quiz score percentage on my vertebrate quiz than I did on my invertebrate quiz.

Each participant was given a link to the survey after completing the vertebrate post-test. Surveys were completed on student iPads and the results were generated into bar graphs and a pie chart.

This concluded the data collection process for the study.

Data Analysis

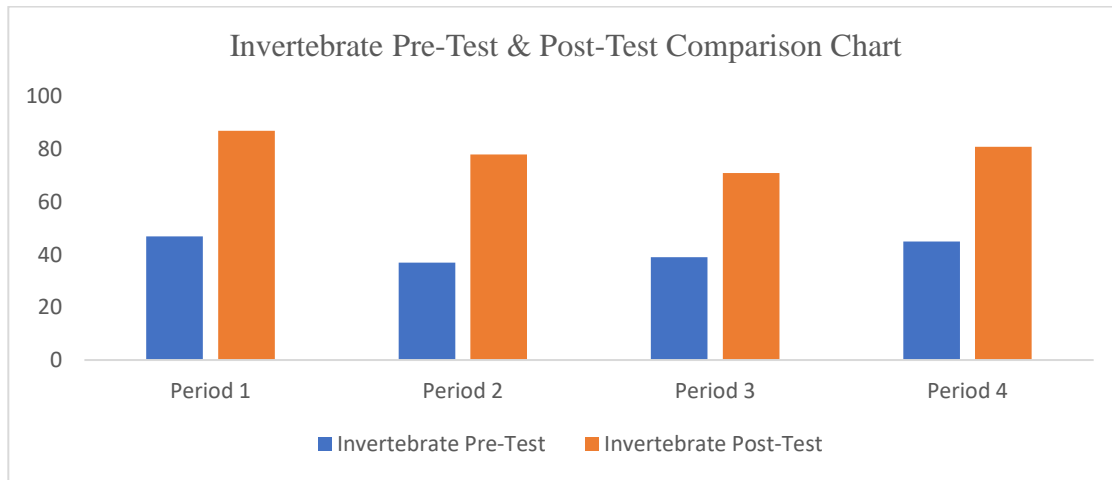
Phase One Findings:

For phase one of the study, a dependent samples T-test was used to determine if student scores improved from the invertebrate pre-test to the post-test. This was chosen to analyze the impact of formative assessment on student learning when no formative assessment framework was embedded into classroom practice by the researcher. First, student scores from the invertebrate pre-test and post-test were organized into data tables for each class period. From here, class period average percentages were determined for each test. Once the mean scores were calculated for each test, the researcher then subtracted the pre-test mean score from the post-test mean score to determine the percent score increase for each class period. The results are represented in Table 1 below and Chart 1 on page 25.

Table 1

Invertebrate Pre-test & Post-Test Scores

| Class Period | Invertebrate Pre-Test Mean % | Invertebrate Post-Test Mean % | % Increase |
|---------------------|-------------------------------------|--------------------------------------|-------------------|
| Period 1 | 47% | 87% | 40% |
| Period 2 | 37% | 78% | 41% |
| Period 3 | 39% | 71% | 32% |
| Period 4 | 45% | 81% | 36% |

Chart 1*Invertebrate Pre-test & Post-Test Comparison Chart*

As the results indicate, each class period had significant score increases from pre-test to post-test. Period two had the greatest increase at 41%, while period three had the least at 32%. This indicates that the researcher's use of formative assessment strategies had a positive impact on student learning, even when a framework was not embedded into classroom practice. These results will be compared to the vertebrate pre-test and post-test results from phase two of the study to analyze any differences.

In addition to the dependent samples T-test data above, Table 2 on page 26 was created to depict the number of student scores within defined percent ranges on the invertebrate post-test for each class period. This was done to show the range of student scores and which percent ranges accounted for the most student scores from each class period. The number in red represents the highest number of student scores within a defined percentage range for each class period.

Table 2*Invertebrate Post-Test Range Scores*

| Class Period | # of student scores between 90-100% | # of student scores between 80-89 | # of student scores between 70-79% | # of student scores below 70% |
|--------------|-------------------------------------|-----------------------------------|------------------------------------|-------------------------------|
| 1 | 7 | 5 | 3 | 1 |
| 2 | 8 | 3 | 3 | 5 |
| 3 | 4 | 3 | 5 | 8 |
| 4 | 5 | 2 | 6 | 1 |

Note. The numbers in red indicate the percent range with the highest amount of student scores per period.

This data indicates that while there was a great percentage increase from pre-test to post-test for each class period, there was still a wide variety of scores. For example, the period three class had more students score below 70% on the post-test than any other percentage range. This indicates some inconsistency in scores and that the informal use of formative assessment by the researcher may have been less effective than the dependent samples T-test suggests. The results from phase one will eventually be compared to the results from phase two to help determine the effectiveness of the independent variable.

Phase Two Quantitative Findings:

For phase two of the study, the same approach was taken as in phase one. To start, a dependent samples T-test was used to determine if student scores improved from the vertebrate pre-test to the post-test. This was chosen to analyze the impact of formative assessment on student learning when William & Thompson's five key strategies framework was embedded into classroom practice by the researcher. First, student scores from the vertebrate pre-test and post-test were organized into data tables for each class period. From here, class period average

percentages were determined for each test. Once the mean scores were calculated for each test, the researcher then subtracted the pre-test mean score from the post-test mean score to determine the percent score increase for each class period. The results are represented in Table 3 and Chart 2 below.

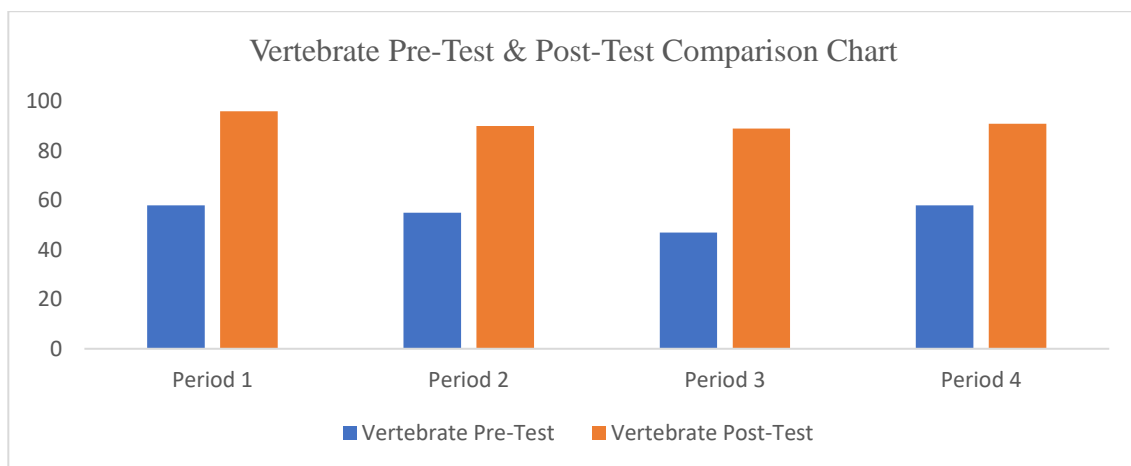
Table 3

Invertebrate Post-Test Range Scores

| Class Period | Vertebrate Pre-Test Mean % | Vertebrate Post-Test Mean % | % Increase |
|--------------|----------------------------|-----------------------------|------------|
| Period 1 | 58% | 96% | 38% |
| Period 2 | 55% | 90% | 35% |
| Period 3 | 47% | 89% | 42% |
| Period 4 | 58% | 91% | 33% |

Chart 2

Vertebrate Pre-test & Post-Test Comparison Chart



As the results indicate, each class period had significant score increases from pre-test to post-test. Period three had the greatest increase at 42%, while period four had the least at 33%. This indicates that the researcher's use of formative assessment strategies when embedding William & Thompson's five key strategies framework had a positive impact on student learning. These results will be compared to the invertebrate pre-test and post-test results from phase one of the study to analyze any differences.

In addition to the dependent samples T-test data above, Table 4 was created to depict the number of student scores within defined percent ranges on the vertebrate post-test for each class period. This was done to show the range of student scores and which percent ranges accounted for the most student scores from each class period. The number in red represents the highest number of student scores within a defined percentage range for each class period.

Table 4

Vertebrate Post-Test Range Scores

| Class Period | # of student scores between 90-100% | # of student scores between 80-89 | # of student scores between 70-79% | # of student scores below 70% |
|--------------|-------------------------------------|-----------------------------------|------------------------------------|-------------------------------|
| 1 | 16 | 0 | 0 | 0 |
| 2 | 12 | 3 | 4 | 0 |
| 3 | 12 | 5 | 2 | 1 |
| 4 | 9 | 4 | 1 | 0 |

Note. The numbers in red indicate the percent range with the highest amount of student scores per period.

This data indicates that in addition to a great percentage increase from pre-test to post-test for each class period, there was also consistency in scores. For example, the period one class had every student score in the 90-100% range. In addition to this, each of the other three class periods also had the most amount of student scores within the 90-100% range. It is also

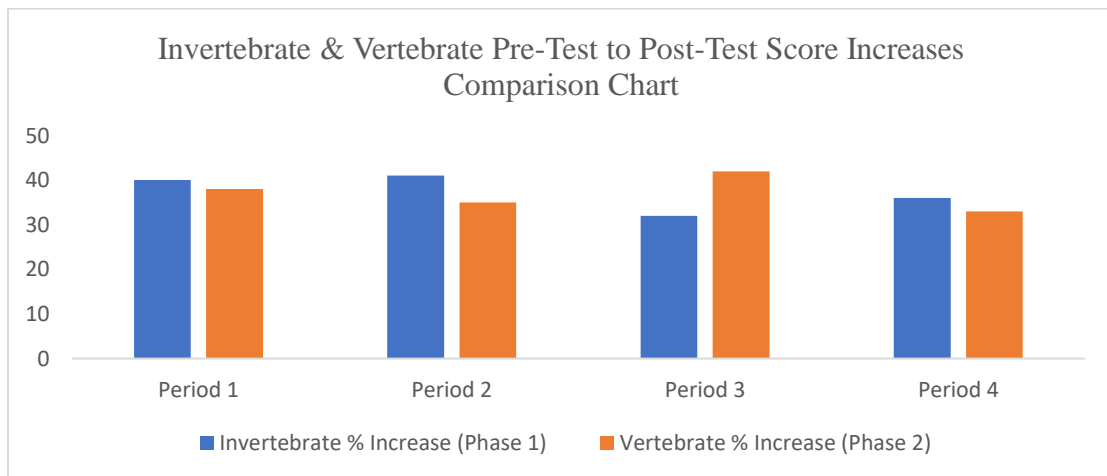
noteworthy that only one student from all four class periods had a post-test score below 70%. In comparison, 15 total students had an invertebrate post-test score below 70%. This indicates more consistency in overall scores on the post-test during phase two of the study. It also provides evidence that the embedding of William & Thompson's five key strategies framework during phase two may have had a greater overall impact on student learning than the informal use of formative assessment during phase one.

Phase One & Two Quantitative Data Comparisons:

In order to better determine the impact of the independent variable on the dependent variable in this study, the data from phases one and two need to be compared. The first set of data that will be compared is the pre-test to post-test score increases. While the results from each phase are similar, the phase one pre-test to post-test scores are slightly better overall. This is represented in the Chart 3 below.

Chart 3

Invertebrate & Vertebrate Pre-test to Post-Test Score Increases Comparison Chart



Note. The invertebrate percent increase represent phase one of the study. The vertebrate percent increase represent phase two of the study.

As can be seen, there was a greater increase in the pre-test to post-test scores in three of the four class periods during phase one. This could imply that the informal use of formative assessment during phase one of the study was more effective than when it was embedded during phase two. While this could be true, it seems too simple and does not give a well-rounded data picture to jump to such conclusions.

A second set of data that will be compared is the number of student scores within defined percent ranges on the post-tests of each phase. This data is represented in the Table 2 and Table 4 below. It is also represented in Charts 4-7 on pages 31 and 32.

Table 2

Invertebrate Post-Test Range Scores

| Class Period | # of student scores between 90-100% | # of student scores between 80-89 | # of student scores between 70-79% | # of student scores below 70% |
|--------------|-------------------------------------|-----------------------------------|------------------------------------|-------------------------------|
| 1 | 7 | 5 | 3 | 1 |
| 2 | 8 | 3 | 3 | 5 |
| 3 | 4 | 3 | 5 | 8 |
| 4 | 5 | 2 | 6 | 1 |

Note. The numbers in red indicate the percent range with the highest amount of student scores per period.

Table 4

Vertebrate Post-Test Range Scores

| Class Period | # of student scores between 90-100% | # of student scores between 80-89 | # of student scores between 70-79% | # of student scores below 70% |
|--------------|-------------------------------------|-----------------------------------|------------------------------------|-------------------------------|
| 1 | 16 | 0 | 0 | 0 |
| 2 | 12 | 3 | 4 | 0 |
| 3 | 12 | 5 | 2 | 1 |
| 4 | 9 | 4 | 1 | 0 |

Note. The numbers in red indicate the percent range with the highest amount of student scores per period.

Chart 4

Period 1 Invertebrate & Vertebrate Post-Test Score Comparisons Chart

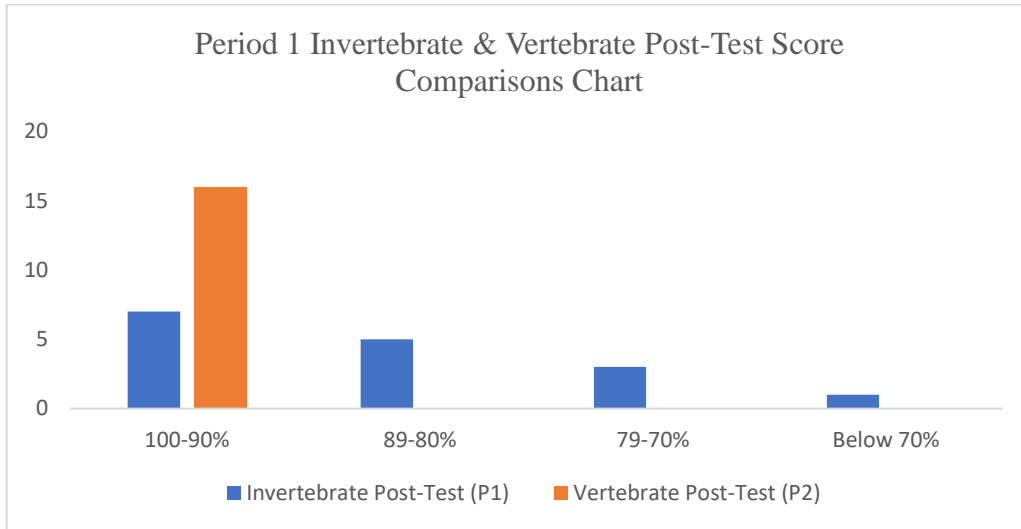


Chart 5

Period 2 Invertebrate & Vertebrate Post-Test Score Comparisons Chart

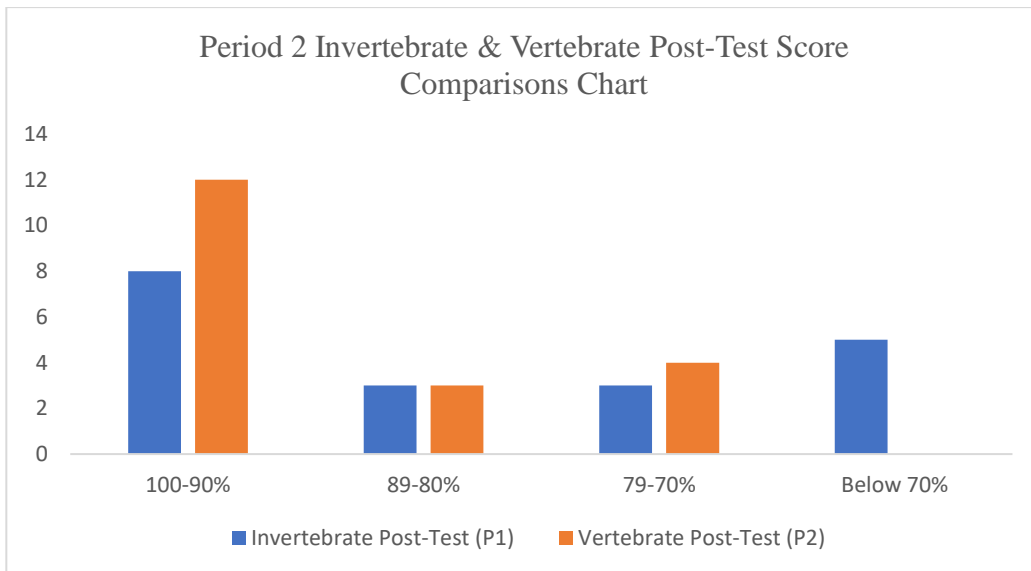


Chart 6

Period 3 Invertebrate & Vertebrate Post-Test Score Comparisons Chart

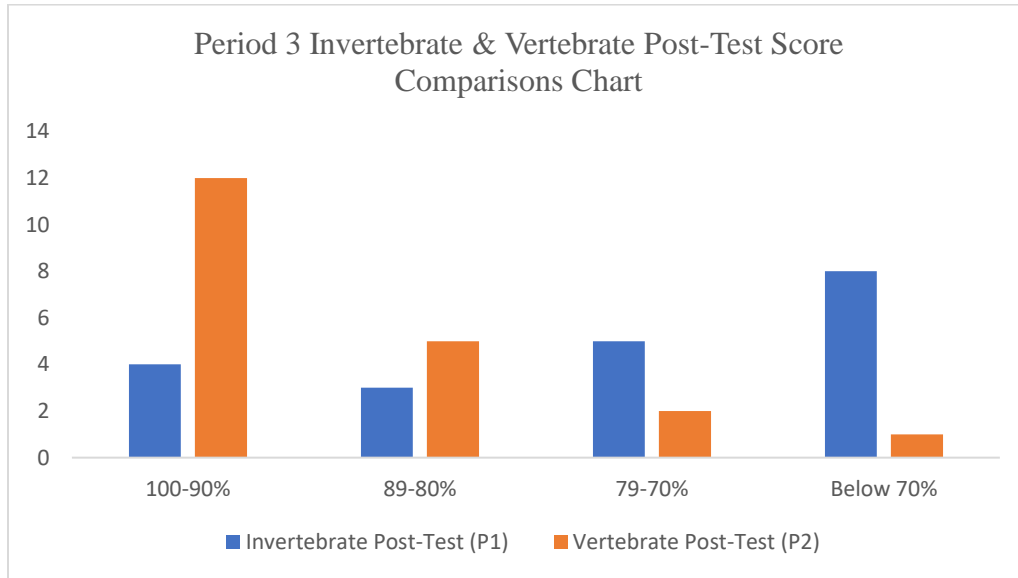
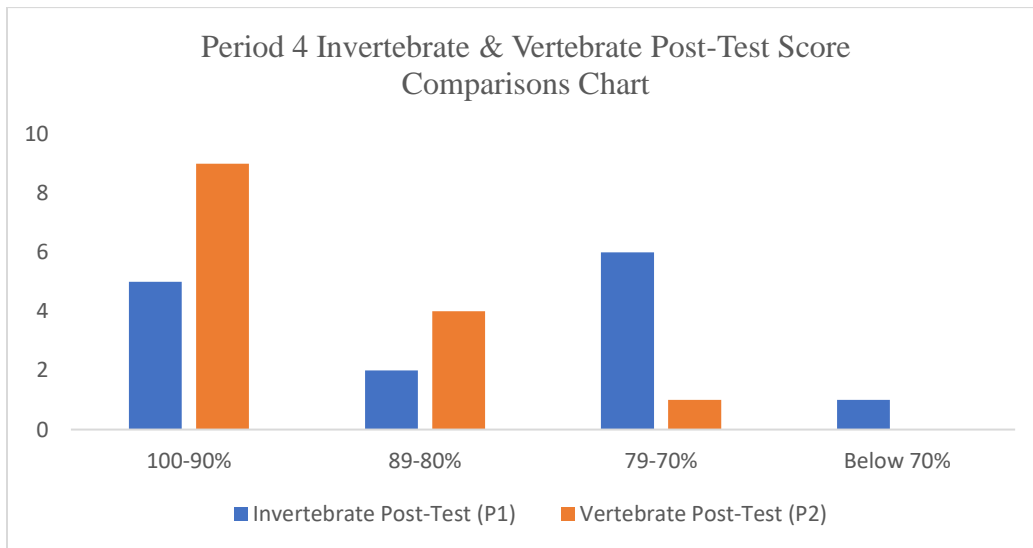


Chart 7

Period 4 Invertebrate & Vertebrate Post-Test Score Comparisons Chart



What becomes clear with this set of data is that student scores on the phase two post-test were higher and more consistent than the phase one post-test scores. For example, the period one class had all 16 students score in the 90-100% range on the post-test during phase two. During the phase one post-test, less than half of these same students scored in the 90-100% range. In addition to this, each of the other three class periods also had most of their students score in the 90-100% range. In comparison, not a single class period during phase one had more than 43% of students score in the 90-100% range. As mentioned earlier, it is also noteworthy that only one student from all four class periods had a post-test score below 70% during phase two. In comparison, 15 total students had a phase one post-test score below 70%.

These data results certainly show that the phase two post-test scores were far more consistent than the phase one scores. It also shows that the phase two post-test scores were much higher overall than the phase one post-test scores. The four graphs above provide a great visual that reinforces this reality. Overall, the results imply that the students received a more effective and equitable education during phase two of the study when William & Thompson's five key strategies framework was embedded into classroom practice.

A third set of data that will be compared will be each phase's measures of central tendency. This set of data is shown in Table 5 and Table 6 on page 34.

Table 5*Vertebrate Post-Test Measures of Central Tendency Chart*

| Class Period | Invertebrate Pre-Test Mean | Invertebrate Post-Test Mean | Post-Test Mode (Whole #s) | Post-Test Range | Post-Test S.D. |
|--------------|----------------------------|-----------------------------|-----------------------------|-----------------|----------------|
| Period 1 | 47% | 87% | Scores in the 10s. (91-99%) | 4.62 | 1.14 |
| Period 2 | 37% | 78% | Scores in the 10s. (91-99%) | 6.08 | 1.94 |
| Period 3 | 39% | 71% | Scores in the 8s. (72-81%) | 7.5 | 2.15 |
| Period 4 | 45% | 81% | Scores in the 8s. (72-81%) | 6.2 | 1.55 |

*Note. Mode scores were calculated by converting partial point test scores into whole numbers.***Table 6***Vertebrate Post-Test Measures of Central Tendency Chart*

| Class Period | Vertebrate Pre-Test Mean | Vertebrate Post-Test Mean | Post-Test Mode (Whole #s) | Post-Test Range | Post-Test S.D. |
|--------------|--------------------------|---------------------------|-----------------------------|-----------------|----------------|
| Period 1 | 58% | 96% | Scores in the 13s. (93-99%) | 1.25 | 0.39 |
| Period 2 | 55% | 90% | Scores in the 13s. (93-99%) | 3.68 | 1.28 |
| Period 3 | 47% | 89% | Scores in the 13s. (93-99%) | 5.65 | 1.33 |
| Period 4 | 58% | 91% | Scores in the 13s. (93-99%) | 3.14 | 1.06 |

Note. Mode scores were calculated by converting partial point test scores into whole numbers.

Mean

Beginning with a comparison of the post-test mean scores, it is clear to see that this favored the phase two results. During phase two, the highest class period mean was 96%. During phase one, the highest class period mean was 87%. During phase two, three of the four periods posted mean scores of at least 90%. During phase one, not a single class period posted a mean score above 87%. During phase two, the lowest class period mean was 89%. During phase one, the lowest class period mean was 71%.

Mode

Mode scores also favored phase two results. Since post-test scores included partial points for both phases, the researcher decided to calculate the mode using only whole numbers. For example, if a score was either 13 or 13.33, both would be counted as scores of 13. This was done to make the mode more realistic with such a large variety of partial scores. During phase two, the mode score was 13 for all four class periods. A score in the 13s represented a percentage score between 93-99% on the post-test. During phase one, only two class periods had mode scores within a 91-99% percent range. The other two period's mode scores of 8, represented percentage scores between 72-81% on the post-test.

Range

Range scores also favored phase two results. During phase two, the lowest post-test range of 1.25 was posted by period one. During phase one, the lowest post-test range was 4.62. During phase two, three class periods had range scores of 3.68 or lower. No class period range scores could match this during phase one. During phase two, the highest range score was 5.65. During phase one, the highest range score was 7.5.

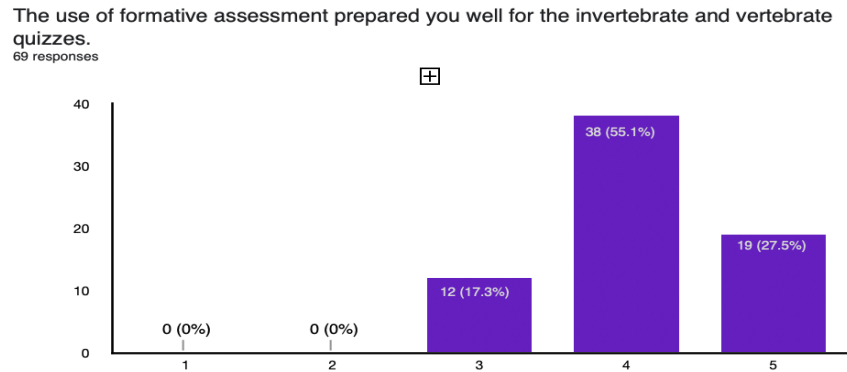
Standard Deviation

The final measurement displayed in the data table is each class period's standard deviation. This measures how dispersed the data is in relation to the mean. A lower standard deviation would mean more scores or data clustered around the mean. This measurement also favored the phase two post-test results. Each class period's standard deviation was lower on the phase two post-test than on the phase one post-test. Additionally, period one only had a standard deviation of 0.39. This is favorable for the phase two results because more student scores were clustered around the higher overall mean scores.

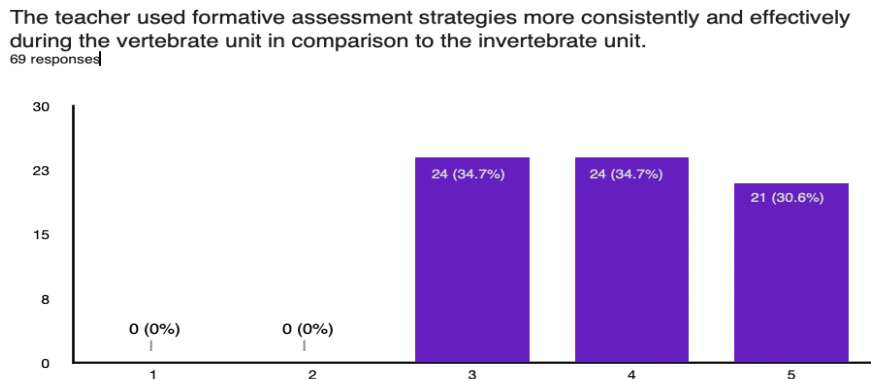
When analyzing each of the measures of central tendency, it is clear to see the data favors the phase two results. Overall post-test scores were both higher and more consistent. This set of data implies that students received a more effective and equitable education during phase two of the study. It also implies that the researcher's use of formative assessment had a greater impact on student learning when it was embedded into classroom practice using William & Thompson's five key strategies framework.

Qualitative Data (Student Survey):

This final set of data was gathered using student surveys. The five-question survey was created using Google Forms and utilized a five-point Likert scale. Each question was designed to use student feedback to help determine any impact formative assessment had on their learning experience. Three of the questions specifically gave students a chance to share their feedback regarding the phase two unit of study. This feedback will be analyzed to help determine if the independent variable of the study had any noticeable impact on student attitude and learning. The results of the survey were generated into Charts 8-12 below.

Chart 8*Student Survey Question #1 Chart*

With this question, 82.6% of students either “agreed” or “strongly agreed” that the use of formative assessment prepared them well for both phases post-tests. 17.3% of students responded neutrally, while 0% of students “disagreed” or “strongly disagreed” with the statement.

Chart 9*Student Survey Question #2 Chart*

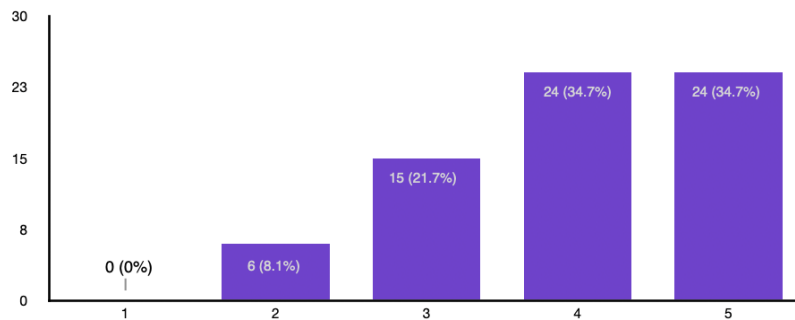
With this question, 65.3% of students either “agreed” or “strongly agreed” that the teacher used formative assessment more consistently and effectively during phase two. 34.7% of students responded neutrally, while 0% of students “disagreed” or “strongly disagreed” with the

statement. This shows that most students did recognize a more consistent and effective use of formative assessment during phase two.

Chart 10

Student Survey Question #3 Chart

I felt more confident and prepared when studying for the vertebrates quiz than I did for the invertebrates quiz
69 responses

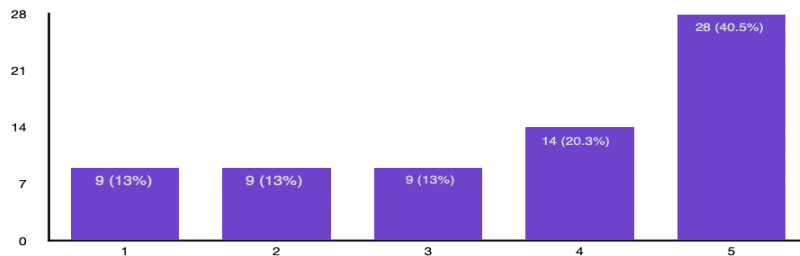


With this question, 69.4% of students either “agreed” or “strongly agreed” that they were more confident and prepared when studying for the phase two post-test. 21.7% of students responded neutrally, while 8.9% of students “disagreed,” and 0% “strongly disagreed.” Students were asked to respond to this statement to see if the use of formative assessment during phase two may have led to greater student confidence than during phase one.

Chart 11

Student Survey Question #4 Chart

I felt more confident while taking the vertebrates quiz than I did while taking the invertebrates quiz
69 responses

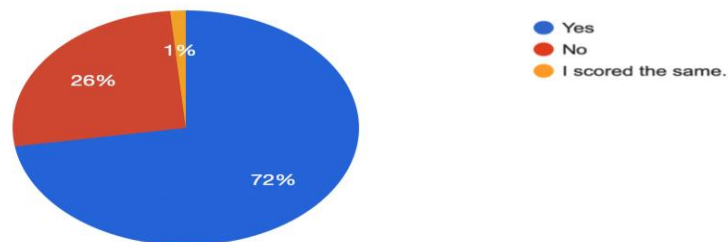


With this question, 60.8% of students either “agreed” or “strongly agreed” that they felt more confident while taking the phase two post-test. Only 13% of students responded neutrally, while 13% of students “disagreed,” and 13% “strongly disagreed.” This indicates that most students felt more confident while taking the phase two test. This may imply that they were better prepared for the test due to the embedding of formative assessment during phase two. With that being said, the fact that 26% of students disagreed with this statement lessens the validity of drawing such conclusions.

Chart 12

Student Survey Question #5 Chart

I had a higher quiz percentage on the vertebrate quiz than the invertebrate quiz.
69 responses



Overall, this set of qualitative data serves to provide a more well-rounded data picture by including student feedback. A few things are clear as a result of the survey. First, most students believed that the use of formative assessment strategies better prepared them for each phases post-test. In fact, not a single student disagreed with this statement. Second, most students recognized and believed that the researcher used more consistent and effective formative assessment strategies during phase two of the study. Third, confidence increased for most students during phase two of the study. Lastly, most students had improved phase two post-test scores. The overall qualitative data picture may also imply that the researcher’s use of formative

assessment during phase two of the study had a greater impact on student learning when it was embedded into classroom practice using William & Thompson's five key strategies framework.

Discussion

The completion of this action research study resulted in a couple major findings. The first major finding was that the use of formative assessment can have a positive impact on student learning. This was confirmed during each phase due to the significant increases in the pre-test to post-test scores. During phase one, the average class period's pre-test to post-test score increased by 37.25%. During phase two, the average class period's pre-test to post-test score increased by 37%. These results supported the use of formative assessment as an effective teaching strategy.

The second major finding was that when William & Thompson's five key strategies framework is embedded into classroom practice, the use of formative assessment can have an increased positive impact on student learning. While this finding was not confirmed by the dependent samples T-test of pre-test to post-test scores, it was supported by each measure of central tendency. For example, three out of four class periods posted mean scores of at least 90% on their post-test during phase two. During phase one, not a single class period posted a mean score above 87% on their post-test. During phase two, the lowest class period mean was 89% on the post-test. During phase one, the lowest class period mean was 71% on the post-test. In addition, the mode, range, and standard deviation also improved during the phase two post-test. With every measure of central tendency favoring the phase two results, it was clear that when William & Thompson's five key strategies framework is embedded into classroom practice, the use of formative assessment can have an increased positive impact on student learning.

Another finding was that students recognized the more consistent and effective use of formative assessment strategies when William & Thompson's five key strategies framework was

embedded into classroom practice. This was supported by the question #2 results on the survey where 65.3% of students either “agreed” or “strongly agreed” that the teacher used formative assessment more consistently and effectively during phase two. In addition, 0% of students disagreed with the statement of the question in any way.

The findings of this study could have a significant impact on the teaching and learning process. As discussed in the literature review, many educators recognize the impact that formative assessment can have on student learning. While this is the case, several researchers also believe that formative assessment is not being used effectively by many teachers. To correct this problem, researchers have hypothesized that educators can embed a high-quality formative assessment framework into classroom practice. The findings of this study support this suggestion since student learning increased at a greater rate when William & Thompson’s framework was embedded into classroom practice than when it was not. Teachers can use these findings to improve their use of formative assessment by embedding a formative assessment framework into classroom practice. This, in turn, could result in improved student learning.

While this study does support embedded formative assessment as an improved formative assessment approach, it does have limitations. For example, the study had a relatively small sample size and duration time. Another limitation was that only one teacher participated in the study. Increased teacher participation would provide more validity to the results of the study. A final limitation was that the independent variable was applied to all students and there was no separate student control group. As a result, the independent variable had to be applied to a different unit of study so that all students received it. This was done for ethical reasons so that some students were not possibly receiving a lower-quality education than others during the study.

Further Study

Based on the results of this study, future research on formative assessment should continue to focus on how to best integrate formative assessment into classroom practice. While it is clear the use of formative assessment can positively impact student learning, there is still far less clarity on how to best implement formative assessment strategies. Future research should continue to focus on formative assessment frameworks and how to effectively embed them into classroom practice. This study focused on Williams & Thompson's five key strategies framework, but there are several other proposed frameworks that can be studied. By focusing on embedded formative assessment, educational researchers can identify successful formative frameworks and integration strategies.

Another area for further study would be formative assessment professional development. With formative assessment research trending towards embedding frameworks and models into classroom practice, teachers will inevitably need to receive training on how to do this. This need for training has been confirmed by recent research studies that have shown that teachers' formative assessment practice can vary greatly despite similar educational backgrounds and environmental contexts (Sathasivam & Daniels, 2016). Studies have also shown that teacher self-efficacy regarding formative assessment practice, can also vary dramatically even amongst teachers within the same school setting (Xiang, et al., 2020). Professional development sessions will need to be created and researchers will need to investigate the impact of professional development on teachers' implementation of formative assessment frameworks. By doing this, researchers will be able to identify effective professional development methods that will help teachers successfully use formative assessment in their classrooms.

Lastly, future research will need to focus on how to successfully integrate formative assessment into the online learning environment. With education continuing to move towards online learning platforms and classrooms, educators will need to adapt and learn how to best use formative assessment in this environment. Digital formative assessment tools, software programs, and learning applications are being created to meet this demand. Specifically, formative assessment software is being created that can help deliver dynamic real-time feedback (Bhagat & Spector, 2019). This ability for technology to provide dynamic real-time feedback, can help students take ownership of their learning like never before (Dalby & Swan, 2019). While this is the case, educators still need to be trained on how to use such software in the online learning environment (Bhagat & Spector, 2019). Educational researchers will need to research the impact of these tools and how they can be used to help teachers embed formative assessment into the online classroom.

Conclusion

The concept and use of formative assessment continue to develop and evolve over time. What began as a simple alternative to traditional summative assessment, has developed into a variety of teaching strategies, frameworks, and theories on classroom practice. Today, the successful implementation of formative assessment into classroom practice seems to be the predominant focus of study. This is a result of recent research showing that many educators are struggling to effectively integrate high-quality formative assessment into classroom practice.

In response, researchers have developed formative assessment frameworks and proposed embedding these frameworks into classroom practice to improve teachers' use of formative assessment. While researchers have touted the embedding of these frameworks into classroom practice as a solution to poor formative assessment practice, there is little empirical evidence to

support these claims. That is where this study found its purpose. The purpose of this study was to investigate if formative assessment had a greater impact on student learning when it was embedded into classroom practice using William & Thompson's five key strategies framework.

Upon completion of this study, the data supports the researcher's hypothesis that embedding Williams and Thompson's five key strategies framework into classroom practice would increase student learning at a greater rate than through the informal use of formative assessment. This was supported by the fact that most of the quantitative data improved during phase two of the study when the researcher embedded William & Thompson's framework into classroom practice. During this portion of the study, student post-test scores significantly increased, and every measure of central tendency improved. The qualitative data also supported this conclusion by conveying that students believed that formative assessment was used more consistently and effectively during phase two of the study. While there were limitations to this study, the results indicate that embedded formative assessment is a promising formative assessment methodology that can improve teacher practice and student learning.

References

- Almuntasheri, S. (2016). Saudi teachers' practices of formative assessment: a qualitative study. *Problems of Education in the 21st Century*, 74, 6–15.
- Andersson, C., & Palm, T. (2017). Characteristics of improved formative assessment practice. *Education Inquiry*, 8(2), 104–122.
- Bhagat, K. K., & Spector, J. M. (2017). Formative Assessment in Complex Problem-Solving Domains: The Emerging Role of Assessment Technologies. *Journal of Educational Technology & Society*, 20(4), 312–317. <http://www.jstor.org/stable/26229226>
- 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.
- Beesley, A. D., Clark, T. F., Dempsey, K., & Tweed, A. (2018). Enhancing formative assessment practice and encouraging middle school mathematics engagement and persistence. *School Science and Mathematics*, 118(1-2), 4–16. <https://doi.org/10.1111/ssm.12255>
- Bernard, P., Dudek-Różycki, K., & Orwat, K. (2019). Integration of inquiry-based instruction with formative assessment: the case of experienced chemistry teachers. *Journal of Baltic Science Education*, 18(2), 184–196. <https://doi.org/10.33225/jbse/19.18.184>
- Brink, M., & Bartz, D. E. (2017). Effective use of formative assessment by high school teachers. *Practical Assessment, Research & Evaluation*, 22(8).
- Dalby, D., & Swan, M. (2019). Using digital technology to enhance formative assessment in mathematics classrooms. *British Journal of Educational Technology*, 50(2), 832–845. <https://doi.org/10.1111/bjet.12606>
- DeLuca, C., Chapman-Chin, A. E. A., LaPointe-McEwan, D., & Klinger, D. A. (2018). Student perspectives on assessment for learning. *Curriculum Journal*, 29(1), 77–94. <https://bera-journals-onlinelibrary-wiley-com.ezproxy.nwciowa.edu/doi/full/10.1080/09585176.2017.1401550?sid=worldcat.org>
- Dini, V., Sevian, H., Caushi, K., & Picón, R. O. (2020). Characterizing the formative assessment enactment of experienced science teachers. *Science Education*, 104(2), 290–325. <https://doi.org/10.1002/sce.21559>
- Hegazy, H., & Barton, G. (2017). Formative assessment in the middle years: a review of literature and alignment with the Guiding Principles for Junior Secondary.

- Hopfenbeck, T. N., & Stobart, G. (2014). *(PDF) assessment and learning: State of the Field Review. - researchgate*. Retrieved July 25, 2022, from https://www.researchgate.net/publication/263654863_Assessment_and_Learning_State_of_the_Field_Review
- Johnson, C. C., Sondergeld, T. A., & Walton, J. B. (2019). A Study of the Implementation of Formative Assessment in Three Large Urban Districts. *American Educational Research Journal*, 56(6), 2408–2438
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Klute, M., Apthorp, H., Harlacher, J., Reale, M., Regional Educational Laboratory Central (National Center for Education Evaluation and Regional Assistance, (Marzano Research, L. (2017). Formative Assessment and Elementary School Student Academic Achievement: A Review of the Evidence. REL 2017-259.
- Menon, P. (2018). Role of assessment conversations in a technology-aided classroom with english language learners: an exploratory study. *Multicultural Education*, 25(2), 42–50. <https://www.proquest.com/docview/2055552600?accountid=28306>
- Naghdipour, B. (2017). Incorporating formative assessment in iranian efl writing: a case study. *Curriculum Journal*, 28(2), 283–299. Article 15
- Pinger, P., Rakoczy, K., Besser, M., & Klieme, E. (2018). Interplay of formative assessment and instructional quality-interactive effects on students' mathematics achievement. *Learning Environments Research*, 21(1), 61–79. <https://doi.org/10.1007/s10984-017-9240-2>
- Sathasivam, R. V., & Daniel, E. G. S. (2016). Tale of two science teachers' formative assessment practices in a similar school environment. *Asia-Pacific Forum on Science Learning and Teaching*, 17(2).
- Schildkamp, K. (2020). *Formative assessment: A systematic review of critical ... - sciencedirect*. Retrieved March 1, 2023, from <https://www.sciencedirect.com/science/article/pii/S0883035520300082>
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné & M. Scriven(Eds.), *Perspectives of curriculum evaluation* (Vol. 1, pp. 39-83). Chicago, IL:Rand McNally.
- William, D., & Thompson, M. (2008) Integrating assessment with instruction: what will it take to make it work? In C. A. Dwyer (Ed.) *The future of assessment: shaping teaching and learning* (pp. 53-82). Mahwah, NJ: Lawrence Erlbaum Associates.

- William, D., & Thompson, M. (2011) Integrating assessment with instruction: what will it take to make it work? In C. A. Dwyer (Ed.) *The future of assessment: shaping teaching and learning* (pp. 53-82). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wools, S., & Veldkamp, B. P. (2019). A general framework for the validation of embedded formative assessment. *Jem*, 56(4), 715–732. <https://doi.org/10.1111/jedm.12234>
- Xiang, X., Yum, S., & Lian, R. (2020). Teachers' self-efficacy and formative assessment of students: moderating role of school goal structure. *Social Behavior and Personality*, 48(6), 1–12. <https://doi.org/10.2224/sbp.9208>