


8-2020

Understanding The Importance Of Formative Assessment Programs In Undergraduate Medical Education

Matthew Gentile
University of New England

Follow this and additional works at: <https://dune.une.edu/theses>

 Part of the Educational Assessment, Evaluation, and Research Commons, Educational Leadership Commons, Higher Education Commons, and the Medical Education Commons

© 2020 Matthew Gentile

Preferred Citation

Gentile, Matthew, "Understanding The Importance Of Formative Assessment Programs In Undergraduate Medical Education" (2020). *All Theses And Dissertations*. 316.
<https://dune.une.edu/theses/316>

This Dissertation is brought to you for free and open access by the Theses and Dissertations at DUNE: DigitalUNE. It has been accepted for inclusion in All Theses And Dissertations by an authorized administrator of DUNE: DigitalUNE. For more information, please contact bkenyon@une.edu.

UNDERSTANDING THE IMPORTANCE OF FORMATIVE ASSESSMENT PROGRAMS IN
UNDERGRADUATE MEDICAL EDUCATION

By

Matthew Gentile

BA (Richard Stockton University) 2005

MA (Richard Stockton University) 2008

A DISSERTATION

Presented to the Affiliated Faculty of

The College of Graduate and Professional Studies at the University of New England

Submitted in Partial Fulfillment of Requirements

For the degree of

Doctor of Education

Portland & Biddeford, Maine

August, 2020

Copyright 2020
By Matthew Gentile

Matthew Gentile
August 2020
Educational Leadership

UNDERSTANDING THE IMPORTANCE OF FORMATIVE ASSESSMENT PROGRAMS IN UNDERGRADUATE MEDICAL EDUCATION

ABSTRACT

The purpose of this quantitative research study was to investigate the impact and relationship formative assessment in undergraduate medical education has on summative assessment performance. Previous research has investigated the important role formative assessment has on the broad education process (Menéndez, Napa, Moreira, & Zambrano, 2019). This research study examined that role in undergraduate medical education. This research study utilized a preexisting de-identified data set consisting of 332 third-year medical student assessment records from five different cohorts. The initial analyses investigated the differences in summative assessment performance for students who completed a formative assessment program during their third year of medical school and those who did not experience the formative assessment program in seven core clerkship disciplines. Further analyses investigated student records for relationships between undergraduate medical education student formative assessment program performance and summative examination performance in the seven core clerkship disciplines. The research study quantitative data analyses evaluated performance differences and relationships among groups of undergraduate medical students exposed to a formative assessment program at the study site. The researcher found multiple statistically significant results. Exposure to the formative assessment program had a statistically significant impact on summative assessment performance. Furthermore, formative assessment performance had a statistically significant

relationship with summative assessment performance. These findings hold importance for students in assessing their knowledge strengths and weaknesses, faculty members in the design of their educational plans, and organizational leadership when making decisions about support for formative assessment programs at their institutions.

Keywords:

Assessment, Formative Assessment, Summative Assessment, Undergraduate Medical Education, NBME® Subject Examination, Formative Assessment Programs, Medical Knowledge Assessment

University of New England

Doctor of Education
Educational Leadership

This dissertation was presented
by

Matthew Gentile

It was presented on
July 24, 2020
and approved by:

William Boozang, EdD, Lead Advisor
University of New England

Debra Welkley, EdD, Secondary Advisor
University of New England

Cheryl Melovitz-Vasan, PT, DPT, PhD, Affiliate Committee Member
Cooper Medical School of Rowan University

ACKNOWLEDGEMENTS

I would like to express the most profound appreciation to my dissertation committee, Drs. William Boozang, Debra Welkley, and Cheryl Melovitz-Vasan. Their encouragement, patience, and support throughout the entire dissertation process was invaluable. I would like to thank Dr. Susan Perlis for her supportive mentorship in the field of medical education. I would like to thank my parents Joseph and Joan, and my brother Joseph for their encouragement throughout my life to always work hard to achieve my goals. Finally, I would like to thank my wonderful wife, Angela, and my children for the support and understanding they showed as I worked to earn a doctoral degree.

TABLE OF CONTENTS

CHAPTER 1	1
INTRODUCTION	1
Problem Statement	3
Purpose of the Study	6
Research Questions	7
Conceptual Framework	8
Assumptions, Limitations, Scope	10
Significance	12
Definition of Terms	13
Conclusion	13
CHAPTER 2	15
LITERATURE REVIEW	15
Problem Statement & Significance	16
Conceptual Framework	19
Personal Interest	19
Topical Research	20
Theoretical Framework	20
Literature Review	22
Educational Assessment	23
Medical Education Assessment	27
Formative Assessment Perception	30
The Role of Leadership in Formative Assessment	32
Conclusion	36
CHAPTER 3	38

METHODOLOGY	38
Research Questions	38
Research Design	39
Hypotheses	43
Data Set	43
Data Collection.....	44
Data Analysis	47
Limitations	49
Internal and External Validity	52
Ethical Issues	53
Conclusion and Summary	55
CHAPTER 4	57
RESULTS	57
Analysis Method	57
Presentation of Results	62
Descriptive Data	64
Hypothesis 1	68
Hypothesis 2	71
Summary	74
CHAPTER 5	75
CONCLUSION	75
Interpretation of Findings.....	77
Hypothesis One Interpretation.....	78
Hypothesis Two Interpretation.....	81
Student Perception.....	85
Faculty Perception.....	85
Administrative Leadership Perception	86

Implications	87
Recommendations for Action.....	89
Recommendations for Further Study	90
Conclusion.....	96
REFERENCES	99
APPENDIX A	111
APPENDIX B.....	114

LIST OF TABLES

Table 1. Hypotheses One and Two Research Design.....	63
Table 2. Descriptive Data by Graduating Year	65
Table 3. Core Clerkship Formative Assessment Program Performance	66
Table 4. Core Clerkship Summative Assessment Performance	68
Table 5. Independent Samples t-tests: Efficacy of Formative Assessment Program	69
Table 6. Strength of Relationship between Formative and Summative Assessment Performance	73

LIST OF DATA SETS IN APPENDIX B

Data Set 1. Family Medicine Clerkship scatter plot.....	114
Data Set 2. Internal Medicine Clerkship scatter plot.....	115
Data Set 3. Neurology Clerkship scatter plot	116
Data Set 4. OB-GYN Clerkship scatter plot.....	117
Data Set 5. Pediatrics Clerkship scatter plot	118
Data Set 6. Psychiatry Clerkship scatter plot	119
Data Set 7. Surgery Clerkship scatter plot.....	120
Data Set 8. Overall average performance scatter plot	121

CHAPTER ONE

INTRODUCTION

Undergraduate medical education institutions work to admit and retain students who can become competent physicians after completing their medical education program (Doukas & Volpe, 2018). A primary goal for undergraduate medical education institutions and students is that upon graduation, students possess a wealth of medical knowledge that allows them to achieve success as a physician in the field of medicine (Farrell, Bourgeois-Law, Buydens, & Regehr, 2019). According to Bass (2008), “the greater a member’s responsibility for attaining a goal, the stronger his or her commitment to the goal” (p. 773). Students and institutions set goals and objectives they must achieve before a student can graduate and become a physician. It is imperative that students have access to feedback about their performance and current knowledge-based strengths and weaknesses while working through the education process as they obtain their degree. Achieving goals is an essential part of the learning process for students. It is imperative that students strive to build a needed knowledge-base while reaching their learning goals. (McConnell, Harms, & Saperson, 2016).

The term *assessment* was defined by the Standards for Educational and Psychological Testing as “any systematic method of obtaining information from tests and other sources, used to draw inferences about characteristics of people, objects, or programs” (AERA, APA, & NCME, 1999, p. 172). There is a movement in the medical education community to share more information about aggregate student performance with individuals and other institutions as a means of feedback about learning goals and gaps in knowledge (Eltorai, 2013). A method of communication about student progress through formative assessment systems allows for the sharing of performance information with the ability for improvements before taking their

summative assessments (Elmahdi, Al-Hattami, & Fawzi, 2018). Formative assessment is vital in undergraduate medical education because of the foundation it builds, leading to later summative assessments. This same feedback is useful for organizational leaders as well because they are gaining insight into their students and their strengths and weaknesses (Raupach & Schuelper, 2018). This same feedback is helpful for institutions to understand better where the gaps in their curriculum exist in preparing medical students to become physicians. According to Leggio and Albritton (2015), “feedback is information provided to improve or optimize performance” (p. 163). Without some standardization to assess progress within the learning process, students, faculty, and organizational leaders do not know how well their medical education program and its students are performing until delivering final summative assessments, at which time negative results are damaging to the student (Gullo et al., 2015).

The focus of this research study was undergraduate medical education student performance from a Northeastern Medical School, given the pseudonym “Northeastern Medical School” for the purposes of this research study. This institution is a public M.D. degree-granting medical school with an enrollment of under 500 students across four years of medical education. This research study focused on third-year medical student performance on their National Board of Medical Examiners (NBME) developed summative subject examinations in seven different core clerkship disciplines encountered during the third year of their medical education. Students complete their summative subject examinations after their third year of medical school during an intensive three-week testing period. During the academic year, students complete formative subject examinations also developed by the NBME after their concentrated six-week block of time in one of the seven core clerkship disciplines. These seven core clerkships included Internal Medicine, Family Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, and

Surgery. The formative assessment tool, titled the “clinical science mastery series,” are practice assessments developed by the NBME and made available for purchase to both individual students and medical schools. Each discipline has its own formative assessment tool. The primary purpose of these tools is to help students prepare for the summative subject examinations (“Comprehensive Self-Assessment Services,” 2019). The worthiness of this research study is evident because of the importance medical education and healthcare communities place on the attainment of medical knowledge outcomes, which were measured by the NBME subject examinations after the third year of medical school at Northeastern Medical School. The summative subject examinations were a major component of each grade in the seven core clerkship disciplines accounting for 35% of each overall clerkship grade. The other parts of each clerkship grade are comprised of clinical assessments related to patient care activities.

The researcher works as an administrator at Northeastern Medical School and investigated strength in the relationship between delivered formative assessments and summative assessment performance to provide support for the use of the formative assessment program in undergraduate medical education. Research evidence showing support for the relationship between the formative assessment program and summative assessment performance would affirm the importance of the program for undergraduate medical education students at Northeastern Medical School, administrators, and the broader undergraduate medical education community with the common goal of student attainment of strong medical knowledge foundation.

Statement of the Problem

Undergraduate medical education students work to build medical knowledge before graduation and need consistent performance feedback related to their academic progress while

building this knowledge. The feedback about their knowledge-base informs them about how well they are achieving their goals and the goals set by the undergraduate medical education institution during the education process (McConnell, Harms, & Saperson, 2016). Medical students have no way of knowing if they are studying in the right manner, reviewing the correct materials, or grasping the high-yield information needed to obtain desired summative assessment outcomes without formative performance feedback leading to their summative assessments (Watling, 2016). Formative assessment bolsters the individual student's understanding of their strengths and weaknesses (Deiglmayr, 2018). By creating a mandatory formative feedback assessment system, students, faculty, and administrators can gain the needed feedback about how prepared students are for summative assessments in medical education and beyond their graduation from the institution.

Formative assessment programs are essential in undergraduate medical education to reinforce learning and help students in their preparation for future career goals (Dolin, Black, Harlen, & Tiberghien, 2018). Formative assessment that does not provide direct feedback to undergraduate medical education students about their performance and current state of knowledge is part of a loosely coupled assessment system. The feedback should inform the student about needed areas of improvement or strengths to build upon in the future. If the feedback fails to reach or impact the student, the feedback will not be helpful (Elmahdi, Al-Hattami, & Fawzi, 2018). The loosely coupled assessment system lacks coordination, regulations are lax, and can result in a lack of self-determination and actualization by medical students about their obtained knowledge (Weick, 1976). The system must be timely and provide direct feedback related to the student's current performance and understanding of a particular area of medical knowledge in the clinical sciences.

Well developed formative assessment systems ensure students are receiving quality information about their current standing within the education they are a part of (Dolin, Black, Harlen, & Tiberghien, 2018). A formalized formative assessment system, coordinated through a student's education institution with centralized monitoring and support, allows students to feel confident in their formative assessment feedback while preparing for high-stakes summative assessments and their future careers. Investigating the importance of instituting required formative feedback programs in undergraduate medical education is needed because of the evidence showing variability in student perceptions about their readiness and how to best prepare for their summative assessments (Kumar, Shah, Maley, Evron, Gyftopoulos, & Miller, 2015). Predicting future performance can be achieved using a gradual process where students recognize personal gaps and seek help to remedy those gaps. Contributions to the field in this area should add an understanding of how beneficial these programs are for students, faculty, learning support staff, and leadership in undergraduate medical education. Recognizing student deficiencies and strengths requires more urgency to keep up with the demand for competent and confident physicians. According to Beaudoin (2012), organizational practices must be malleable to keep up with changing environmental factors surrounding the institution. Kotter (2012) describes the importance of organizational urgency concerning needed changes. Because medical student summative performance and attainment of medical knowledge is crucial to future career success and the impact well-trained future physicians have on society, instituting a robust mandatory formative feedback system is an urgent change that needs implementation throughout undergraduate medical education. This research study sought to examine a robust school-wide formative feedback system and determine its influence on improving student performance outcomes and informing school leadership about needed program improvements. The

overarching problem driving this research study is that greater support is needed for formative assessment programming in undergraduate medical education to support preparation for summative assessments and longterm attainment of medical knowledge.

Purpose of the Study

The purpose of this research study is to investigate the impact and relationship formative assessment in undergraduate medical education has on summative assessment performance. Previous research has investigated the important role formative assessment has on the broad education process (Menéndez, Napa, Moreira, & Zambrano, 2019). This research study affirmed the role of formative assessment in undergraduate medical education, as its use to support learning throughout the education spectrum is critical (Deiglmayr, 2018). This research study utilized a preexisting de-identified data set consisting of 332 third-year medical student assessment records from Northeastern Medical School. The data set encompassed five different cohorts of students from the graduating classes of 2016 to 2020. The planned analyses investigated the differences in summative subject examination performance for students who completed the formative assessment program during their third year of medical school and those who did not experience the formative assessment program in the seven core clerkship disciplines completed at Northeastern Medical School.

Further analyses investigated student records for evidence of relationships between undergraduate medical education student formative assessment program performance and summative examination performance in the seven core clerkship disciplines. Significant statistical findings showing meaningful differences between students who did and did not complete formative assessment and significant statistical relationships between performance on formative assessments and summative examinations support the need for formative assessment

programs. These findings are most relevant to Northeastern Medical School, but also for the broader medical education community. These results aid undergraduate medical education administrators, faculty, and students in building formative assessment programs into undergraduate medical education student assessment systems. The results also help in student preparation and performance on summative assessments and building medical knowledge for their careers.

Research Questions

The overarching research question for this research study is, to what extent do medical knowledge-based formative assessments impact and relate to medical knowledge-based summative assessment performance in undergraduate medical education? The following research sub-questions additionally guide this research study:

1. What effect do medical knowledge formative assessment programs have on individual clerkship assessment performance in undergraduate medical education?
2. What is the extent of the relationship between formative assessment performance and summative assessment performance in undergraduate medical education?

As part of this research study, multiple hypotheses were tested for further investigation of the research topic.

Hypotheses

H₀: Third-year medical students who completed formative assessments during their third year of undergraduate medical education score significantly higher on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education.

H₁: Third-year medical students who completed formative assessments during their third year of undergraduate medical education exhibit no difference in their performance on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education.

H₀: There is a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations.

H₁: There is no significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations.

Conceptual Framework

A conceptual framework is an essential portion of the research process because it provides the foundation through which the research is built (Weaver-Hightower, 2014). According to Ravitch and Riggan (2017), “a conceptual framework is an argument about why the topic one wishes to study matters, and why the means proposed to study it are appropriate and rigorous” (p. 5). One critical aspect of implementing a mandatory formative assessment system is the information provided about student preparedness to complete their summative assessments. Formative assessment related to attained medical knowledge is a critical factor for improving the outcomes for medical students and their schools.

A mandatory formative assessment system requires student assessment in a testing environment that simulates summative examinations. This type of program can provide progress tracking and preparation for critical summative assessments. According to Chang and Wimmers

(2016), regular formative assessment aids in the monitoring process of student achievement.

Konopasek, Norcini, and Krupat (2016) found that formative assessment in medical education cannot be a series of single events, but a natural process that students embrace over time or in a series of events.

The application of a theoretical framework is imperative during the development and design phase of a research study to fully understand the direction of the project (Mills & Bettis, 2015). The theoretical framework for this research study focuses on two published theories that drove the research project and provided the lens used to view this research. The first is aspects of adult learning theory. Knowles (1968) described andragogy as adult learning, which provides the foundation for Adult Learning Theory. Adult learners are more independent and find motivation through internal factors where they focus on attaining only the information they believe they need to know. This type of learning strategy is problematic for medical students because so much of their learning is self-directed. Students have no way of knowing all the medical information needed to be successful future physicians without receiving feedback on their performance, knowledge, and skills. In this planned research study, the relationships between the formative assessment program and summative assessment performance were analyzed. If data shows the formative assessment program had a statistically significant impact on summative examination performance and a statistically significant relationship between the formative and summative assessments, there is support for third-year undergraduate medical education formative assessment programs surrounding preparation for summative medical knowledge assessments and future careers. This supports students using valuable feedback gained through participation in the formative assessment program to improve practice. As described in andragogy (Knowles, 1968), they will not be left to search for answers about readiness on their own.

Second are components of transformational learning theory developed by Mezirow (1978), who described the importance of student reflection on their own experiences to validate and reformulate their understanding of obtained knowledge and learning. The education process should lead to rational choices by students to create reformulated learning plans to achieve learning goals after receiving feedback about their knowledge. The application of this theoretical perspective calls upon students to question themselves and critically assess if they are meeting the desired levels of competency related to their learning goals and objectives. This theory is closely associated with medical student formative assessment because the learning process in medical education is contingent upon the student to recognize their deficiencies and continue to strive for further knowledge acquisition throughout their academic and professional careers.

Assumptions, Limitations, and Scope

There are underlying assumptions made during the creation of a research study (Leedy & Ormrod, 2010). The researcher operated under the assumption that a formative assessment program in undergraduate medical education is beneficial to all students because of the feedback about performance and knowledge the formative assessment program provides over a defined period. The assumption that formative assessments are beneficial to the educational process has support from the literature review that follows and from researchers' work in the medical education assessment field (Downing, & Yudkowsky, 2009; Konopasek, Norcini, and Krupat, 2016). More specifically, is the assumption that there is a relationship between how students perform on their formative clinical science mastery assessments and how they perform on their summative subject examinations in the seven different core clerkship disciplines individually and overall. Another assumption is that students are striving to achieve to the best of their ability on each formative assessment even though it does not count towards their final grade. The

researcher also assumes that all students want to achieve the highest score possible on their summative assessments at the end of the academic year when the summative assessments occur. A final assumption is that, if this research shows that the formative assessment program has an impact and is related to summative assessment performance, it was valuable to Northeastern Medical School in making decisions about continuing the formative assessment program and to the greater undergraduate medical school community in adopting a similar formative assessment program.

In many instances, limitations are inherent to a study and out of the researcher's control, leading to an inability to eliminate those limitations within the framework of the study (Leedy & Ormrod, 2010). A potential limitation of this research study is related to bias about the structure of the formative assessment program by the researcher who works at the institution. A thorough literature review was included in Chapter Two that supports the rationale used to carry out the research study. This literature review built the structure supporting the importance of a formal formative assessment program in undergraduate medical education to help prepare undergraduate medical education students for their high-stakes summative assessments and future careers as physicians. This review showed the need for this research regardless of the researcher's opinions on the topic. Another limitation of this research study is the limited amount of data the researcher could include in the research. Northeastern Medical School has been in existence for less than one decade, so the amount of data within the school's data repository on this topic is limited. Another limitation is that the data for this research is limited to one medical school.

The scope of the research study specifically focused on undergraduate medical education student performance on the formative clinical science mastery series assessments and their relationship with summative Subject Examination assessments at Northeastern Medical School,

in the United States, after the third year of their medical education. Student records were requested for the graduating classes of 2016 through 2020 to investigate a formative assessment program developed to help prepare undergraduate medical education students to take their summative examination in the seven core clerkship disciplines. These summative assessments comprise 35% of the final clerkship grades and are considered high-stakes examinations.

Significance

Finding justification surrounding a formal formative assessment system at the undergraduate medical education level required extensive time, effort, and resources to complete. Investigating the relationship between the formative assessment performance and the summative assessment outcomes allows organizational leaders to promote the significance of the formative assessment system throughout the organization, help students to understand their strengths and weaknesses in preparing for their summative assessments and careers, and help undergraduate medical education students and institutions attain their goals within the medical education community. The investigation and promotion of the formative assessment system is critical because organizational leaders dedicate financial and staff resources to their programs. Statistically significant findings surrounding the impact and relationship between the formative assessment program and summative assessment outcomes could also be relevant to the greater medical education community where a common goal is to ensure students attain needed medical knowledge and achieve their professional goals. Since all American medical students must obtain medical knowledge to graduate from medical school in order to practice medicine, the analysis of formative assessment programming and its relationship to summative assessment performance has potential transferability to the entire undergraduate medical education community.

Definition of Terms

- *Assessment*- The process by which achievement is measured in relation to some course of study (Ferris & O’Flynn, 2015).
- *Formative Assessment*- Evaluation of a student that does not count towards their final grade in an educational activity in any way but has the purpose of informing the student about their performance strengths and weaknesses (Downing & Yudkowsky, 2009).
- *Summative Assessment*- Evaluation of a student that counts towards their final grade in an educational activity (Downing & Yudkowsky, 2009).
- *NBME® clinical science mastery series*- these are assessment tools available to students and institutions for purchase with similar content specifications as the clinical science subject examinations. Assessment tools are available in multiple areas, and when completed, provide a student with extensive feedback reports about their performance (“Clinical Science Mastery Series,” 2019).
- *NBME® Subject Examination*- The National Board of Medical Examiners develops clinical science subject examinations for most core disciplines. Most accredited medical schools use these assessments to assess basic and clinical knowledge of medical students after their core clerkships (Ryan et al., 2017).

Conclusion

Creating a formal formative feedback system in which all undergraduate medical education students participate could enhance student performance on summative assessments, inform faculty and administrative members’ understanding of how prepared students are for critical summative assessments, and strengthen faculty’s knowledge of student attainment of goal achievement. Without a standardized formative assessment system, students, faculty, staff, and

administrators have minimal knowledge about the degree to which their educational program is impacting students and preparing them for high-stakes summative assessments and professional practice. Identification of student and program deficiencies while in the educational process is imperative for improvement and helping students to achieve their professional goals and the goals of the organization for them as graduates and future physicians. Investigating the relationship between the formative assessment program and summative assessment outcomes is a step towards improving the quality of all undergraduate medical education institutions helping their students achieve their goals. Chapter Two provides a detailed review of the current literature surrounding the research study topic and conceptual framework. Future chapters provide an explanation related to the methodology used to analyze the research questions, analysis of findings, and a discussion about the interpretation of those results.

CHAPTER TWO

LITERATURE REVIEW

The purpose of this research study was to investigate the impact and relationship formative assessment in undergraduate medical education has on summative assessment performance. The role of formative assessment in learning throughout the education spectrum is critical because of the role it plays in learning outcomes (Deiglmayr, 2018). A wealth of information related to formative assessment in education exists within the current literature. However, there are gaps in the research about the critical role formative assessment plays in undergraduate medical education and the need for formalized formative assessment systems. The following comprehensive literature review of formative assessment presents the topic from multiple educational perspectives with a primary focus on investigating the role of formative assessment in supporting the learning of undergraduate medical students and the institutions that support them reaching their medical knowledge goals.

Without formative assessment feedback, learners operate without clear direction on how to improve their performance, achieve predetermined goals in the learning environment, and meet the objectives connected to their learning goals and those of their educational institution (Downing & Yudkowsky, 2009). According to Norcini et al. (2018), “effective formative assessment is typically low stakes, often informal and opportunistic by nature, and is intended to stimulate learning” (p. 2). Often undervalued in medical education, formative assessment is an essential component of the educational process. According to Popham (2013), there is a consensus on the general meaning of formative assessment, but no overarching accepted definition for precisely what the term encompasses. In one of the foundational studies in the field, Black and Wiliam (2017) concluded that formative assessment is essential because it

provides insight for all students but, most importantly, underachieving students. The authors also discussed the need for clarification between educators and learners about boundaries between formative and summative assessments in their landmark study. The authors found that educational gains are possible by different routes of learning, but the presence of formative assessment provides a higher likelihood for the most exceptional student achievement.

Medical school leaders can embrace the implementation of formative assessments by allocating budgetary support and enthusiastically promoting the creation of formative assessment systems within their educational program. For this research, the terms formative assessment, formative feedback, and formative assessment programming appear interchangeably. The operational definition for these terms is a set of nationally normed practice examinations, available to undergraduate medical education students that align with the material and presentation of their summative clinical science discipline-specific subject examinations. These formative assessments gauge the degree of obtained medical knowledge a student has learned in their first three years in undergraduate medical education.

Problem Statement and Significance

Undergraduate medical education students work to build medical knowledge before graduation and need consistent performance feedback related to their academic progress while building this knowledge. Feedback about their knowledge-base must inform students about how well they are achieving their goals and the goals of their undergraduate medical education institution during the education process (McConnell, Harms, & Saperson, 2016). Medical students have no way of knowing if they are studying in the right manner, reviewing the correct materials, or grasping the high-yield information needed to obtain desired summative assessment outcomes without formative performance feedback leading into their summative assessments.

Formative assessment bolsters the individual student's understanding of their strengths and weaknesses (Deiglmayr, 2018). By creating a mandatory formative feedback assessment system, students, faculty, and administrators can gain the needed feedback about how prepared students are for summative assessments in medical education and beyond their graduation from the institution.

Formative assessment programs are essential in undergraduate medical education to reinforce learning and help students in their preparation for future career goals (Dolin, Black, Harlen, & Tiberghien, 2018). Medical school leader support for use of a formal formative assessment system to the entire medical community stresses the vital role that formative assessment has in the undergraduate medical education assessment process. Formative assessment that does not provide direct feedback to undergraduate medical education students about their performance and current state of knowledge is part of a loosely coupled assessment system. This type of assessment system has a lack of coordination, lax regulations, and can result in a lack of self-determination and actualization by medical students about their obtained knowledge (Weick, 1976). The formative assessment system must be timely and provide direct feedback related to the student's current performance and understanding of a particular clinical subject area.

Planned valid formative assessment systems ensure students are receiving quality information about their current standing within the education they are a part of (Dolin, Black, Harlen, & Tiberghien, 2018). A formalized formative assessment system, coordinated through their education institution with centralized monitoring and support, can allow students to feel confident in their formative assessment feedback while preparing for high-stakes summative assessments and their future careers. Investigating the importance of instituting required

formative feedback programs in undergraduate medical education is needed because of the evidence showing variability in student perceptions about their readiness and how to best prepare for their summative assessments (Kumar, Shah, Maley, Evron, Gyftopoulos, & Miller, 2015).

Predicting future performance can be achieved through a gradual process where students recognize personal gaps in knowledge and seek help to remedy those gaps. Contributions to the field in this area should add an understanding of how beneficial these programs are for students, faculty, learning support staff, and leadership in undergraduate medical education. Recognizing student deficiencies and strengths requires more urgency to keep up with the growing need for competent and confident physicians practicing medicine.

According to Beaudoin (2012), organizational practices must be malleable to keep up with changing environmental factors surrounding the institution. Kotter (2012) describes the importance of organizational urgency concerning needed changes. Because medical student summative performance is so high-stakes, attainment of medical knowledge is crucial to future career success. The impact well-trained future physicians have on society is undeniable, thus instituting a robust mandatory formative feedback system is an urgent change that needs implementation throughout undergraduate medical education (Gonsalves & Zaidi, 2016). This research study aimed to document the impact robust school-wide formative feedback systems has on improving student performance outcomes and informing school leadership about needed program improvements. The overarching problem driving this research study is that further evidence is needed to show how formative assessment programming improves undergraduate medical education preparation for summative assessments related to the attainment of medical knowledge.

Conceptual Framework

A conceptual framework is an essential portion of the research process because it provides the foundation through which the research is built (Weaver-Hightower, 2014). According to Ravitch and Riggan (2017), “a conceptual framework is an argument about why the topic one wishes to study matters, and why the means proposed to study it are appropriate and rigorous” (p. 5). Conceptual frameworks encompass support from the personal interests of the researcher, topical research area, and theoretical frameworks (Ravitch & Riggan, 2017).

Personal Interest

Personal interest in research is imperative because it fuels the researcher’s drive and work processes to reach their eventual research goals (Ravitch & Riggan, 2017). There is a combination of curiosity, excitement, and determination that goes along with personal interest in research that is a critical component of building the conceptual framework. Excitement related to this topic for the researcher surrounds the potential to investigate and share an accessible system to all undergraduate medical education institutions that could be beneficial to the students and organization in their quest for achievement and building medical knowledge. As a determined administrative leader finding evidence to support the formative assessment system is imperative to continue receiving budgetary and leadership support for a system that is currently in place at the institution. Interest surrounds this research from an administrative standpoint because positive outcomes would support the need to require formative assessment systems in undergraduate medical education as preparation for summative assessment within other organizations while gaining support at the school where the research is to be conducted.

Topical Research

Topical research describes the available information surrounding the topic of interest derived from previously published scholarly works on the thematic research area (Ravitch & Riggan, 2017). This study's research work encompasses the importance of formative assessment in undergraduate medical education. A mandatory formative assessment system requires student assessment in a testing environment that simulates summative examinations. This type of program can provide progress tracking and preparation for critical summative assessments.

According to Chang and Wimmers (2016), regular formative assessment aids in the monitoring process of student achievement. The formative assessment must inform the medical student of their performance, occur promptly, and be documented for students to revisit. Spaced out formative assessment can be supplemented to avoid overassessment and to leave students with the perception that only summative assessments are essential. Konopasek, Norcini, and Krupat (2016) found that formative assessment in medical education cannot be a series of single events. Instead, it must be a natural process that students embrace. Attainment of medical knowledge through a formative assessment system can be vital to the medical education process. Embracing and encountering formative assessments can be equally important. Both personal interest and the research topic are the starting points for understanding the theoretical framework component of the conceptual framework.

Theoretical Framework

The application of a theoretical framework is imperative during the development and design phase of a research study to fully understand the direction of the project (Mills & Bettis, 2015). Grasping a clear understanding of what theoretical frameworks mean can be challenging. Anafara and Mertz (2015) defined theoretical frameworks as “any empirical or quasi-empirical

theory of social and/or psychological processes, at a variety of levels (e.g., grand, midrange, explanatory), that can be applied to the understanding of phenomena” (p. 15). The theoretical framework provides the research consumer with the necessary background and scope needed to ascertain the rationale for conducting the research study.

The theoretical framework for this research study focuses on two published theories that drive the research project and provide the lens used to view this research. First, are aspects of adult learning theory. Knowles (1968) described andragogy as adult learning, which provides the foundation for Adult Learning Theory. Adult learners are more independent and find motivation through internal factors where they focus on attaining the information they believe they need to know. Adult learners also base their need for knowledge acquisition on their own experiences and problems they have already encountered in life. Inherently, adult learners seek to understand why they are learning something before taking steps to acquire new knowledge or skill (Knowles, 1989). This type of learning strategy is problematic for medical students because so much of their learning is self-directed, and students have no way of knowing all the medical information they need to know to be successful future physicians without receiving feedback on their performance, knowledge, and skill.

Merriam (2001) discusses self-directed learning as a process where learners develop their own goals in various aspects of working through their acquisition of knowledge and skill. Both the field of medicine and society require physicians to have a ubiquitous knowledge-base to perform their duties. Since a great deal of medical education is based on the tenets of adult learning and is self-directed, formative feedback about knowledge acquisition and performance is crucial to the process. Without some standardization to assess progress within the learning process, students, faculty, and leaders have no sense of medical education system outcomes.

These crucial outcomes are unknown until delivering final summative assessments, at which time negative results are damaging to the student, the organization, and society.

Second, are components of transformational learning theory developed by Mezirow (1978), who described the importance of student reflection on their own experiences to validate and reformulate their understanding of obtained knowledge and learning. The education process should lead to rational choices by students to create reformulated learning plans to achieve learning goals after receiving feedback about their knowledge. The application of this theoretical perspective calls upon students to question themselves and critically assess if they are meeting the desired levels of competency related to their learning goals and objectives. This theory is closely associated with medical student formative assessment because the learning process in medical education identifies deficiencies in students who continue to strive for further knowledge acquisition throughout their academic and professional careers. These two theories provide the foundation and lens for this research study, stressing the importance of administratively-led structured formative assessment programs in undergraduate medical education with their connectedness to independent learning and ongoing assessment of one's ability while continuously striving to improve their knowledge, skill, and expertise.

Literature Review

The concept of formative educational assessment was first discussed by Scriven (1967), who commented on the diverse role of formative versus summative assessment in education and the importance formative assessment played in fostering a developmental continuum for student learning. Formative assessment as a topic was analyzed from multiple educational perspectives in this research endeavor. The focus of the research project was placed on the impact a formative assessment program has on summative assessment performance and the strength of the

relationship between the two forms of assessment. Bloom (1968) concluded that formative assessment was most important to the educational process by highlighting potential areas of improvement for both learners and educators. The author viewed the acquisition of formative feedback information as the most reliable method to improve ability and knowledge while students are progressing through their education. Without performance feedback, learners progress without clear direction on how to improve their performance and further their education. Often overlooked in the education process, performance feedback is an essential component of the educational process.

According to Popham (2013), there is a consensus on the general meaning of formative assessment, but no overarching accepted definition for precisely what the term encompasses. In one of the seminal works related to formative assessment, Black and Wiliam (2017) concluded that formative assessment is essential because it provides insight for underachieving students. The authors also discussed the need for communicated clarification of boundaries of formative and summative assessment to the learner and educator. The authors profess that educational gains are possible by different routes, but the presence of formative assessment provides a greater likelihood for the highest student achievement (Black & Wiliam, 2017).

Educational Assessment

Educational performance assessments usually take the form of an instrument or tool used to ascertain how well learners are achieving goals and objectives planned and delivered through their educational experiences. Educational formative assessment is used to inform learners about their performance and does not count toward the calculation of a final grade (Yorke, 2013). The information discovered through formative assessment is also useful for faculty and

administrators because they can ascertain gaps in learning and knowledge before the delivery of a summative assessment.

Nitko and Brookhart (2011) concluded that formative assessment is imperative in providing learners with the needed information to improve their future performance. Formal assessment structuring can help predict future performance, guide faculty to bolster teaching in required areas of learner improvement, inform the students about potential gaps in needed knowledge, track the progression of education, and assist educational leadership in resource devotion while providing student support in all relevant areas. Without an understanding of current academic standing, learners may find it challenging to set goals to increase their knowledge and skill in subsequent educational sessions. Black and Wiliam (2009) presented a paradigm related to formative educational assessment. The authors identify feedback development processes and show the need for feedback to be dynamic on presentation, guided, gradual, and controlled for the learner to get the best experience following the feedback delivery. Saunders (2014) concluded,

. . . that there was a need to focus on improving formative assessment and ensuring that the students understood the occurrence and value of this feedback. It is recognized that for feedback to be most useful it should provide students with information that they can directly use to improve subsequent work. (p. 170)

Further, Shute and Kim (2016) discussed the importance of formative educational assessment and described feedback about performance that is not outwardly obvious to the student. This type of feedback represents stealth methods of feedback, where students receive constructive feedback in a less traditional casual manner. Because of the relaxed nature of this form of feedback, the authors concluded that students are more receptive to the feedback process because

they can process the information at their own pace. This type of formative assessment is not commonplace but can be helpful because the student feels they are gaining knowledge and skills gradually and on their own while feedback processes are occurring in the background of their educational experience.

Miller (2009) found that computer-based formative assessments provide students with meaningful feedback about their performance. Because students utilize technology to complete the assessment, many formative assessment mediums can provide highly efficient and detailed feedback reports to students immediately following the completion of the assessment. Demir (2018) concluded that technological-based assessment creates a significant impact on participation in the assessment process because of levels of comfort and ease of use. O'Leary, Scully, Karakolidis, and Pitsia (2018) concluded that technological advancements related to educational assessment are continuously evolving and make the education process more dynamic.

Educational summative assessments evaluate student performance and often measure the degree to which students have achieved the educational goals presented for the educational activity at all levels of education (Downing & Yudkowsky, 2009). Summative assessments equate to some form of a final score or grade that informs the student about final performance (Nitko & Brookhart, 2011). The results from summative assessments are outcomes data that help determine student readiness, goal acquisition, and knowledge. In many instances, summative assessments serve as a competency rating to decide if a learner can move forward to the next level in their education and result in some form of a grade.

Maintaining a balance between providing a student with a final measured performance report, and their acknowledgment of the information about their strengths and weaknesses

related to that performance, is imperative in their growth of knowledge. According to Burke (2010), the summative assessment must be the result of a balanced approach to informing students about their performance while also allowing room for improvement. If the pendulum is weighted too heavily in either the formative or summative assessment tools, learners may find it challenging to obtain the most enriching educational experience.

Chappuis (2014) presented information related to the need for well-planned assessments to understand the meaning and depth of student performance across the educational spectrum. The creation of a positive assessment environment where students are encouraged to take the feedback information and positively move forward with future performance is vital for their success. The author concluded that students must have the opportunity to improve following the delivery of feedback. Improvement serves as the central goal of delivering educational, formative assessment (Chappuis, 2014). If the process gets carried out as planned, educators should provide their educational product to learners within the scope of an informative session. Learners work to acquire knowledge and skill within the delivered educational course. Educators offer some formative assessment tools following the provided educational experience to decipher the acquisition of knowledge and expertise, which learners obtained through the given educational course. Educators then provide feedback to learners related to their formative assessment performance. Learners then continue with their educational process, reinforcing their knowledge and skill through further learning. Repeat formative assessment is possible at some point in the process, but the summative assessment is the true culmination in the educational process to measure if students obtained the required amount of knowledge and skill within the predetermined educational process (Buelin, Ernst, Clark, Kelly, & DeLuca, 2019). This

discussion highlights the importance of further investigation surrounding the impact of a defined formative assessment program and its relationship to summative assessment performance.

Medical Education Assessment

Formative assessment is essential in medical education as it is in other forms of education. According to Moore (2018), when discussing medical education assessment, “formative assessment is feedback and guidance provided continuously throughout courses to help students understand how they are progressing toward accomplishing the goals and objectives of the course and what they need to do to continue progressing” (p. 52). Formative assessment provides learners with vital information to improve their summative assessment performance. After undergraduate medical education, medical students become medical doctors with responsibility for patient care. The medical knowledge obtained while in medical school must support physicians as they encounter patient care and medical practice. As Moore (2018) concluded, the medical education formative assessment process must be continuous throughout the learning process for students as they continue toward the eventual goal of becoming a practicing physician.

The current standard in medical education is to provide students with a roadmap for success and hope to meet them at the finish line as a successful medical school graduate, well-trained to enter the world of clinical practice (Moore, 2018). According to Leggio and Albritton (2015), “feedback is information provided to improve or optimize performance” (p. 163). In most professional arenas and levels of education, performance feedback is imperative because it creates a critical understanding of strengths and weaknesses. In medical school, formative feedback is required for students to gauge their medical knowledge, current clinical ability, areas needing improvement, and how to improve performance with further training.

Formative assessment is a requirement by the Liaison Committee on Medical Education (LCME) that oversees and accredits all United States' and Canadian medical schools. According to the LCME *Scope and Purpose of Accreditation*, "LCME accreditation is a voluntary, peer-reviewed process of quality assurance that determines whether the medical education program meets established standards. This process also fosters institutional and programmatic improvement" (2019). Medical education institutions must strive for the primary objective to prepare medical students to become highly functioning, competent future physicians. The LCME standards serve to ensure the attainment of this goal is in place and achievable. While working towards achieving stated objectives, students must experience multiple forms of formative assessment throughout their medical education careers.

Although formative assessment is mandatory, according to the LCME, no prescribed systems exist related to delivery style and methods in undergraduate medical education. The formative assessment must be relative to the educational experience and included in the educational plan at all levels of education (Rushton, 2005). Without a specified program, students struggle to benchmark where they stand in their medical knowledge acquisition and what they need to achieve for successful summative assessment outcomes and future career achievement.

Students must be made aware of the available formative assessment practices at the beginning of their educational experience so that they can anticipate the points of feedback they will receive about their performance. According to Gruppan, Ten Cate, Lingard, Teunissen, and Kogan (2018), "in particular, the formative uses of assessment data in providing feedback to learners need to be linked closely to the setting and time of the performance" (p. S18). Linkage of formative assessments to summative assessment outcomes and knowledge acquisition is the

result of formative assessment programs that guide students in achieving their individual goals and the goals of the institution. Konopasek, Norcini, and Krupat (2016) concluded that formative assessment systems are most useful to students when the system is a clearly defined process that connects performance events and relays information to the student. The authors of the research study further state that organizing and delivering such a system is difficult because of the undergraduate medical education culture focused on summative assessment grades, but formative assessment system creation is imperative because of the educational needs of medical students.

Students need formative feedback to improve performance but cannot rely entirely on direct educator feedback due to difficulties in standardization in the quality of the input (Konopasek, Norcini, & Krupat, 2016). The nature of validated formative assessment tools that gauge current standing relative to educational goals is the cornerstone of a formative assessment program. Self-reflection is also necessary for continuous improvement (Sadler, 1989). Palmen, Vorstenbosch, Tanck, and Kooloos (2015) found that the more frequently students sought and participated in formative assessments, the stronger their summative assessment performance was later during their educational experience. Mandatory formative assessment systems enhance student understanding of preparedness for summative assessments, inform faculty and administrative members' knowledge of potential student performance outcomes, strengthen faculty knowledge of student attainment of learning objectives, and enlighten institution leadership's questions about students achieving defined program objectives before graduation (Palmen, Vorstenbosch, Tanck, & Kooloos, 2015). Without a standardized formative assessment system, students, faculty, staff, and administrators have difficulty understanding the degree to which the educational program is impacting students and preparing them for summative

assessments and professional practice. Deficiencies in the educational process need identification for improvement and achievement of shared organizational and student goals.

Summative assessment in medical education often equates to some form of a final score or assigned grade based on student performance. This type of assessment is imperative to medical schools because it serves as the assurance to society that a student is prepared to go forward into the medical profession (Pangaro, 2015). According to Downing and Yudkowsky (2009), summative assessment should roll up to some form of final statement about the student related to their ability in a defined area of medicine. Without this information, society cannot be certain of the student's preparation for the next stage in their training or practice of medicine. Brenner, Bird, and Willey (2017) found that standardized curricular formative assessments can predict later summative examination performance, but licensure examination represents the results of a single assessment resulting in one score. The acquisition of medical knowledge throughout undergraduate medical education is the foundation of the medical knowledge a student utilizes for their entire professional career.

Formative Assessment Perception

Perceptions of medical education assessments vary among students and faculty. In their 2017 study that surveyed student and faculty perceptions about formative feedback, Mulliner and Tucker found that the time it takes to deliver, the style in which it is provided, and the ability of faculty member to provide feedback were all found to be essential to student reception and use of feedback. If students are unable to accept or understand formative feedback about their current ability and knowledge, the entire process becomes jaded. Many faculty members struggle with the concept of required formative assessments in medical education. Many support a less formal feedback process. Gardner-Gletty (2002) found that significant work was needed to develop

faculty on the importance of formative assessment in higher education through departmental cultural shifts from previously held notions about the low importance of formative assessment.

According to Harrison et al. (2016), in general, medical students want greater control over the feedback process and need a more structured feedback system than is currently commonplace in the field. These findings directly relate to the previously discussed theoretical framework. As stated in Knowles' (1968) and Mezirow's (1978) theories, student independence, while obtaining information and reflection to validate their independently obtained information, is a key factor in their acceptance of feedback. As introduced by Harrison et al. (2016), student control over the structured feedback is also imperative in conjunction with their independence and validation. Langendyk (2006) found low achieving students could not correctly assess their performance difficulties after completing formative and summative assessments. Kibble (2007) stated that the more frequently students participated in formative assessments, the better their performance on later summative assessments were. While the authors did not report finding that students needed to perform at a high level on the formative assessments, the participation process itself aided in the later summative performance. Deeley (2017) found divergent results related to student perceptions of technology systems used to deliver, house, and report formative assessment performance. The inconclusive results of this study bring the usefulness of the tools into question because they did not meet the needs of all students in the study and potentially negatively impacted the formative assessment goals and outcomes.

At some point in the educational process, a medical educator most likely needs to switch roles from the purveyor of feedback to the final rater of performance and attainment of educational goals (Yorke, 2003). It is during this transition that the critical nature of formative assessment becomes evident to the educator because of the need to rate student performance,

following the previous delivery of performance feedback. Hill (2014) discussed educator perceptions of formative assessment and concluded that faculty members overwhelmingly reported feeling like they received adequate faculty development relative to delivering formative feedback and designing needed formative assessments. Although the analysis dealt with self-report data only, the significant findings indicate a strong understanding of formative assessment among faculty in the population studied. Bok et al. (2013) found that providing formative assessment was not easy for all faculty because of student perceptions of the process and their own curriculum delivery needs. Careful attention is needed for the development of formative feedback systems around faculty skills and support. When the faculty leaders/members used formative assessment correctly, it helped students to identify learning objectives and to improve their study strategies (Al Kadri, Al-Moamary, Magzoub, Roberts, & van der Vleuten, 2011). Close (2017) found that faculty perceptions of formative feedback have a direct impact on the quality of the feedback presented to students. If faculty members viewed the formative assessment process favorably, they were more likely to provide a quality formative assessment to students.

The Role of Leadership in Formative Assessment

Multiple factors impact student attainment of knowledge outside the walls of the educational institution and control of educators surrounding gaps in learning ability and variations in the foundations of the greater educational environment (Guskey, 2005). Designing and assuring implementation of a formative assessment program must be championed by medical education administrative leadership to ensure continuous delivery regardless of student and faculty perceptions. Although the implementation of the formative assessment systems may be a change to an accepted culture, it is imperative for leadership members to gain support for the

deployment because of the potential importance of the system. When discussing the role of educational institutions on assessment practices, Hersh and Keeling (2013) concluded that “assessment must reflect the institution’s collective commitment to the cumulative nature of higher learning and the understanding that assessment done well promotes learning” (p. 12). The organization must commit to learning and providing an assessment to promote education and attainment of organizational goals for learners.

Kotter (2012) described the importance of organizational urgency about needed changes. Although a formative assessment program could be costly, if the evidence supports a need for the implementation, an organization would need to act urgently to support such a system. Bolman and Deal (as cited in Gallos, 2006) note the importance that organizational leaders play in supporting and sustaining change within their organization. Support from leaders must be a consistent effort that includes financial, staffing, and space support for a robust formative assessment system in this change initiative. Without support from leadership, the planned action will not succeed in attaining the important goals related to the proposed program.

Further, Bass (2008) discussed the relationship between leadership and organizational change and presented the importance of esteem about the leadership role. This topic broadly addresses the way other individuals view the person with leadership status. This dynamic between leaders and those they lead is imperative for fostering successful organizational change.

In some situations, medical education leaders face conflict related to the proposed implementation of the formative assessment system from students and faculty who do not agree with the nature of the newly introduced system. Conflict can be challenging in any leadership situation. According to Shetach (2012), “When two parties are communicating, and a conflict occurs, as long as both sides continue to concentrate on the issue, they are headed toward a

satisfactory solution” (p. 26). There are many types of organizational leaders identified in the literature who may contribute to beneficial outcomes while leading their organizations through important organizational change (Bass, 2008). The transformational leader embodies many of the qualities that lead to successful introduction and execution of major organizational change, like the implementation of new programs, processes, and procedures. The transformational leader strives to achieve organizational success through change related to their goals. They must also attempt to motivate followers to adopt the same ethos they espouse as the path to organizational success. Even the charismatic, transformational leader faces conflict and resistance to change along their way. According to Bass (2008), if individuals in leadership roles are not looked at as being correctly assigned to their position as a leader or obtaining requisite knowledge and skill, the conflict undoubtedly occurs between the leader and the group. The biggest challenge related to this style of leadership is gaining acceptance from followers. It can be hard to judge someone’s moral code from their words alone. Followers need to see the transformational leader in action over time to understand their conviction, skill, and ethical character are genuinely dedicated to the betterment of the organization. Medical education curriculum and assessment require more innovative leaders to comply with future changes to medical education (Benjamin, Benjamin, Benjamin, & Selfridge, 2018). Organizational leaders at the educational institution must take strides to implement a formative feedback system while gaining the support of the educational community with the process to ensure students have the best professional opportunities following undergraduate medical education.

Palmer and Devitt (2018) analyzed 259 medical students over two years and found that including components of formative assessment into the curricular design has a positive impact on student performance and engagement in the educational process. Marion and Gonzales (2014)

stated that culture “is a phenomenon that encompasses every element of organizational life” (p. 259). The leader must shift the perception to believe in the need for the formative assessment system for successful implementation. Bass (2008) found that transformational leaders can guide followers to follow their plans through charisma and including followers in the process. Active engagement of students and faculty is imperative to the creation of the formative feedback system.

An embodiment of transformational leadership qualities among medical education leaders ensures the successful implementation of a formative feedback system. Bass (2008) concluded that the transformational leader would have the most significant positive impact on organizational change due to the very foundation of their leadership style and attributes. The environment where the transformational-idealized influence leader is most successful is where the charismatic individual has gained the esteem of those they lead; their moral character shows full investment in followers and organizational success concurrently, and they have a stable moral code, based on their conduct.

According to Marion and Gonzales (2014), an idealized influence type of transformational leadership is hinged on the need for a leader to possess charisma. They further discussed the charismatic leader as having inborn-extraordinary abilities to elicit change from their followers. The successful transformational leader must also embody a robust moral code while putting the needs of the organization ahead of all else. The individuals who respond most productively to this style of leadership are those who see some likeness between themselves and the transformational leader, feel respected by the leader, and believe that the leader can produce some form of success identifiable by the group (Bass, 2008).

According to Bass (2008), “the greater a member’s responsibility for attaining a goal, the stronger his or her commitment to the goal” (p. 773). Medical education leadership tends to become mired in the importance of individual performance markers without maintaining a focus on big picture related to continual student progress. The role of a medical education leader entails overseeing all medical student performance throughout their educational journey. The operational implementation of the formative assessment system is imperative for student success. From an organizational perspective, this type of formative monitoring system allows for school leadership to review student performance in defined program objectives, while also learning about gaps in delivering and assessing curricular content related to those areas (Konopasek, Norcini, & Krupat, 2016). The medical education and healthcare communities benefit from a system like this because it aids institutions and individuals in reaching their educational goals by helping students to improve performance before summative assessments are delivered (Downing & Yudkowsky, 2009).

Conclusion

The reviewed literature supported the impact and relationship formative assessment systems have on summative assessment performance in undergraduate medical education (Konopasek, Norcini, & Krupat, 2016). Further analyses examining implications and connections between undergraduate medical education formative assessment and later summative assessment needs extensive investigation. Little to no research currently exists in the literature investigating knowledge-based formative assessment systems and the impact they have on later summative assessment performance in undergraduate medical education. The results of this study can help guide medical school leaders in implementing needed formative assessment systems and

improve student and faculty perceptions about the importance of formative assessments on their summative assessment performance and medical knowledge attainment.

Education thought-leaders Black and Wiliam (2017) found that formative assessment is crucial in the education process at any level. The authors presented a paradigm for formative assessment, which included multifaceted approaches to delivering the assessment. According to the LCME (2019), accreditation standards require the delivery of formative assessment in undergraduate medical education, but the accreditors are vague in their definition of formative assessment and the degree of needed inclusion.

The operational implementation of the formative assessment system is imperative for student success. Formative assessments must be woven into medical education curricula because a formal plan to monitor student achievement of objectives is needed to ensure student performance at the highest capacity possible (Downing, & Yudkowsky, 2009; Konopasek, Norcini, and Krupat, 2016). From an organizational perspective, this type of formative monitoring system allows for school leadership to review how well students are performing in achieving the defined program objectives, while also learning about gaps in delivering and assessing curricular content related to those areas. The medical education and healthcare communities benefit from a system like this because it aids institutions and individuals in reaching their educational goals by helping students to improve performance before summative assessment delivery, and the transition from student to medical doctor occurs.

CHAPTER THREE

METHODOLOGY

The purpose of this chapter is to define the rationale of the research study; clearly explain the research questions; discuss the utilized research design; provide hypothesis statements related to the predicted findings; define the research study data set; describe the validated tools used in the research study; define data collection methods, and discuss data analysis strategies. More specifically, the purpose of this program evaluation research study, utilizing both comparative and correlational analyses to investigate archival student performance data, was to explore how a formative medical knowledge-based examination program impacts and relates to performance on undergraduate medical student summative medical knowledge-based examinations after their third year of medical school. This approach allows for a deeper understanding of the impact and relationship formative assessment programs can have on future summative assessment performance and acquisition of medical knowledge.

Research Questions

The overarching research question for this research study is to what extent do medical knowledge-based formative assessments impact and relate to medical knowledge-based summative assessment performance in undergraduate medical education? The following research sub-questions additionally guided this study:

1. What effect do medical knowledge formative assessment programs have on individual clerkship summative assessment performance in undergraduate medical education?
2. If formative assessment programs have a statistically significant impact on summative assessment performance, what is the extent of the relationship between formative

assessment performance and summative assessment performance in undergraduate medical education?

Research Design

This research is a quantitative study, and the research design is a program evaluation that investigated differences and relationships among cohorts of third-year undergraduate medical students through archival data review. The author requested institutional review board permission from both the University of New England, and the research study site for the research plan, analysis of existing student performance assessment data stored within the research study site's data repository. Both institutional review board applications clearly stated that the intended research posed no risk to humans as the research would focus on pre-existing data. The author received institutional review board approval for this research study from both institutions (see Appendix A).

The independent variables in this research study included performance scores on seven distinct formative medical knowledge-based examinations while looking at their relationship with summative subject examination scores of the same content area during the third year of an undergraduate medical education curriculum. The dependent variables were scores on summative assessments that resulted in students' discipline-specific clinical science subject examination scores. According to Mohr (1990), significance testing is the most appropriate way to investigate differences among groups who have and have not encountered some form of an independent variable. The author investigated the differences among groups of students who had and had not completed formative assessments through independent samples *t*-tests.

According to Yan, Su, and World Scientific (2009), correlational analysis allows researchers to take a fundamentally sound computational approach to view relationships between

variables. This research study investigation analyzed the relationship between formative assessment performance and summative assessment performance for students who encountered both assessment activities through correlational analyses. According to Creswell (2015), “in correlational research designs, investigators use the correlation statistical test to describe and measure the degree of association (or relationship) between two or more variables or sets of scores” (p. 339). Analyzing the degree of association between the formative and summative assessment performance in this research study was an imperative component for gaining a better understanding of the relationship between the two types of assessments have with each other in undergraduate medical education at the research study institution.

Both the formative and summative examination tools analyzed in this research study were developed by the National Board of Medical Examiners (NBME) and made available to undergraduate medical education students at institutions across the United States and Canada. Each formative assessment tool has a corresponding summative assessment tool in the same clinical discipline or clerkship within the research study site’s undergraduate medical education curriculum. The formative assessment tool, titled the “clinical science mastery series,” are validated assessments developed by the NBME and made available for purchase to both individual students and medical schools. The assessments included multiple-choice questions with content matching the available clinical science subject examinations. The primary purpose of these tools is to help students prepare for their summative clinical science discipline-specific subject examinations (“Comprehensive Self-Assessment Services,” 2019).

Little research exists related to formative assessment programs in undergraduate medical education, but the practical application of formative assessment tools in undergraduate medical education does exist in the literature. Morrison, Smith, Ross, Butler, and Smith (2016) concluded

that the NBME clinical science mastery series formative assessment, used in this research study, was an excellent tool in providing students with an understanding of their current medical knowledge before completion of summative assessments. These findings are a component of why formative assessment is so critical. The present research study called for the practical application of a formative assessment program throughout the educational process. The formative assessment program was used by third-year undergraduate medical students to improve medical knowledge and performance by providing the learner with feedback about their knowledge related to specific clinical disciplines. More recently, Minor, Stumbar, and Bonnin (2019) published findings related to a Family Medicine Clerkship and cited the importance of the NBME clinical science mastery series formative assessment tool, which is the same formative assessment tool used in the current research study being conducted at the Northeastern Medical School. Minor et al. (2019) concluded that the NBME clinical science series formative assessment tools are useful instruments for the Family Medicine field of study. The findings derived from this study were limited to the application of the formative assessment tool in one specific clinic discipline. Although the Family Medicine research project lends support to the validity of the formative assessment tool, it does not refer to the study of an overarching formative assessment program. The current research study at the Northeastern Medical School focused on the need for an overarching formative assessment program utilizing the NBME clinical science mastery tools.

The National Board of Medical Examiners also designed the summative assessments used in the seven different core clerkship disciplines examined in this research study, which students encounter at the conclusion of their third year of undergraduate medical education at the research study site. The students complete their summative subject examinations after their third

year of medical school during an intensive three-week testing period before entering into their fourth year of medical school. The exams are spaced out over three weeks and do not occur on consecutive days of the week. The order of the examinations remains the same every year. During the academic year, as part of their regular undergraduate medical education assessment program, students complete the formative NBME clinical science mastery series assessments described above, after their concentrated six-week block of time in one of the seven core clerkship disciplines. These clerkships included Family Medicine, Internal Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, and Surgery. The formative assessments were made available during the 2017-2018 and 2018-2019 academic years to research study site students in the graduating classes of 2019 and 2020. Before that time, the graduating classes of 2016, 2017, and 2018 completed the summative subject examinations in the seven core clerkships but did not have access to the formative assessment tools. The graduating classes of 2019 and 2020 summative subject examination performance was compared to the summative subject examination performance of the graduating classes of 2016, 2017, and 2018. The classes of 2016, 2017, and 2018 did not complete the formative assessment program, while the graduating classes of 2019 and 2020 did complete formative assessments. All formative and summative assessment performance occurred during the undergraduate medical students' third year of medical school in the academic year before their anticipated fourth year of medical school. Further analysis looked at the relationships between formative and summative assessment performance in the respective clerkships for the classes of 2019 and 2020, who completed both the formative and summative assessments.

Hypotheses

H₀: Third-year medical students who completed formative assessments during their third year of undergraduate medical education score significantly higher on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education.

H₁: Third-year medical students who completed formative assessments during their third year of undergraduate medical education exhibit no difference in their performance on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education.

H₀: There is a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations.

H₁: There is no significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations.

Data Set

The existing assessment performance records for this research study encompassed medical student assessment records on the seven formative NBME clinical science mastery series examinations and the seven summative NBME clinical science subject examinations. Records came from the Northeastern Medical School's five cohorts of third-year undergraduate medical students with anticipated graduation dates of 2016 to 2020. Northeastern Medical

School is a public medical school in the Mid-Atlantic section of the United States. The school has an enrollment of under 500 students and is affiliated with a nearby clinical institution and parent university. The data export came from the school's Progress IQ® assessment data warehouse. The system houses all assessment data for students at the medical school and has an export function to create different data reports using Microsoft® Excel. All student performance metrics were stored in this system as part of the school's standard operating procedures. Data was exported from the Progress IQ® system, de-identified by an independent staff member in the Office of Medical Education at the medical school, and placed into a password-protected Microsoft® Excel spreadsheet for this research study. The de-identified data set included 332 student records for analysis. Students in the data set encompassed five different cohorts, including graduating classes of 2016, 2017, 2018, 2019, & 2020. Full approval to access the de-identified records was obtained from the University of New England, and the research study site institutional review boards before the analysis of the archival data began. During the timeframe of the research study, Northeastern Medical School increased enrollment progressively in all five cohorts analyzed in this research study as part of the institution's strategic plan to increase class size. The increased enrollment caused the number of students in each cohort in this research study to be numerically different. Earlier graduating classes were admitted with small class sizes, making the comparison of three cohorts of students with two cohorts of students valid because the clustered group sizes were relatively close to one another (177 students vs. 155 students).

Data Collection

Following the successful approval of the research proposal by the researcher's dissertation committee, the study site's IRB, and the University of New England's IRB, the data for this research was de-identified by an independent staff member from the research study site's

Office of Medical Education. When the researcher generated the research study site institutional review board applications from the University of New England and the research study site, a notification was sent to the independent staff member to prepare the de-identified data set. The independent staff member had no ties to the research study, and they provided the de-identified data set for the project once the research study site institutional review board approved the research application. The de-identification process is essential for multiple reasons. This process ensures the integrity of the data throughout the research process for the researcher, the research study site, and any affiliated institutions. This process also allows data analysis while protecting the identity of the students' assessment information in this research study (Kayaalp, 2018). No identifiable information was available to the researcher. The researcher was the sole individual to analyze and interpret the data, which remained on a password-protected external hard drive in the researcher's locked office, within a locked cabinet throughout the research study. No other person had access to the password, office key, and cabinet key where the data was protected.

At the Northeastern Medical School, the Office of Medical Education purchased seven different formative NBME clinical science mastery series examinations for the medical students during their third year of medical school, starting with the classes of 2019 and 2020. The third-year curriculum at Northeastern Medical School included seven different six-week core clerkships: Family Medicine, Internal Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, and Surgery. During each block, students were assigned to different clinical offices and departments and asked to accomplish a multitude of objectives for each clerkship. For all seven clerkships, medical knowledge related to that discipline was an objective that must be achieved by the students. On the final day of their six-week clerkship, in all seven blocks, students were brought to the medical school to complete the formative NBME clinical science

mastery series assessment for the clerkship they recently completed. Students completed the formative assessments under proctored conditions to match the testing environment they encounter with later summative examinations.

All formative NBME clinical science mastery series examinations included 50 multiple-choice questions. Scores range from 0 to 100 on all examinations. After students completed their formative assessment, each student received a feedback report through the testing system immediately after completing the assessment detailing categorical performance, areas of strength, and areas of weakness related to their medical knowledge in that particular discipline. The formative assessment feedback report also provides students with a review of all questions on their assessment, along with an explanation about each question's correct answer. Students received instructions from clerkship faculty and medical school administrative leadership to use their formative assessment feedback reports to fill gaps in medical knowledge discovered during their formative assessments.

Following the formative assessment delivery, students began their next clerkship. The students then went through their next six-week clerkship and completed the same steps after the next block with the formative assessment tools and feedback report. This was repeated until students at the research study site completed all seven core clinical clerkships. After the third year curriculum concluded, but before entering their 4th year of undergraduate medical education, students were brought back to the medical school for three weeks of testing. Students completed seven core clinical clerkship summative NBME subject examination assessments made up of 90 to 110 multiple-choice questions. Scores ranged from 0 to 100 on all examinations. Each of these assessments was a must-pass component of the study site's curriculum, and their final score accounted for 35% of their overall clerkship grade. Once the

medical-knowledge assessments were scored by the NBME, students were awarded grades for each clerkship and informed if they could progress into their fourth and final year of undergraduate medical education. Students also received an extensive feedback report about their summative assessment performance following the distribution of grades in each clerkship.

All formative and summative assessments included in the data set were completed at the research study site by third-year undergraduate medical students. The outcomes from all assessments were stored in the Progress IQ® data repository at the research study site as part of their medical education standard operating procedures. Scores and feedback reports were made available to the Office of Medical Education and the student after completing the assessments. Once scores were delivered to the Office of Medical Education at the research study site, they were stored in the Progress IQ® system.

Data Analysis

The organization of data in a clear, concise manner allows the researcher to ensure possible replication and future communication of findings (Sestoft, 2011). The following data organization process allowed the researcher to ensure analysis of all variables and utilization for future replication and sharing of findings. Once the University of New England and the research institution study site institutional review board applications requesting the data set were approved, the data was organized, de-identified by the research study site independent staff member, and hand-delivered to the researcher with the needed password for access. The data set was saved onto a password-protected external hard drive using a Microsoft® Excel spreadsheet. The researcher used IBM® SPSS Version 26 on their computer and transferred the data into IBM® SPSS Version 26 for further data analyses. SPSS is a statistical software program used to analyze data in the social science fields with descriptive, relational, and grouping variable

analysis (Levesque, 2007). The researcher used IBM® SPSS because it is their preferred system because of prior training they received on the program when learning to conduct statistical analysis. IBM® SPSS Version 26 is available for use at the researcher's institution for internal and external research purposes by faculty, staff, and students. For all hypotheses testing, a confidence level of .05 was used to determine significant results.

Descriptive characteristics for each formative and summative assessment were calculated to investigate measures of central tendency, including mean and standard deviation. The researcher then conducted multiple statistical tests to analyze the data.

Hypothesis one predicted that third-year medical students who completed formative assessments during their third year of undergraduate medical education score significantly higher on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education. Independent samples *t*-tests were calculated to test hypothesis one looking for differences in summative assessment performance for students who did and did not complete the formative assessment program for each clerkship and overall average performance. The researcher tested the differences among groups using the independent samples *t*-test function in IBM® SPSS Version 26. Sawilowsky and Hillman (1992) discussed the validity of using the independent samples *t*-test as a method to investigate differences among groups, which was a major component of this study's research questions. Investigating the differences among students who did and did not complete the formative assessment program produced results to test the hypotheses investigating statistically significant differences between students who did and did not complete formative assessments and their summative assessment performance.

Hypothesis two predicted there is a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations. The researcher used IBM® SPSS Version 26 to test hypothesis two to investigate the Pearson's r correlations between formative and summative assessments in the same clerkship disciplines and overall average performance. The researcher investigated the significance and strength of relationships between formative and summative assessments within the individual clerkships and in average performance. Benesty, Chen, Huang, and Cohen (2009), stated that the Pearson's r correlation provides research consumers with information about the relationship between two independent variables in the areas of statistical significance and strength of the relationships between those variables. Pearson's r was used in this research study to investigate the relationship between formative and summative assessment performance in each discipline-specific clinical clerkship. Results from the independent samples t -tests and Pearson's r correlation tests are provided in Chapter 4.

Limitations

Limitations are a part of all studies and cannot be eliminated by researchers. Instead, the limitation is recognized and accounted for by the researcher within the study (Leedy & Ormrod, 2010). This research study utilized data from students enrolled in one undergraduate medical school in the Northeast part of the United States. Student records were requested for the graduating classes of 2016 through 2020 to investigate a formative assessment program developed to help prepare undergraduate medical education students to take their summative clinical science subject examinations and build medical knowledge. One potential limitation in this research study was that the data obtained in the data set and findings presented in later

chapters are not generalizable to all other medical schools in the United States. The data was collected at one research study site, which represents one medical school. Different medical schools will have stronger or weaker students based on admissions, location, history, and available scholarships. This is a potential confounding factor in the analysis of the study data when trying to compare results on the summative NBME subject examinations at different medical schools. Furthermore, even though the NBME subject examinations are used by many medical schools in the United States, the formative assessment tools developed by the NBME are not used by the same number of schools. This situation makes any broad comparisons of the research study site formative assessment program to all other institutions impossible (Ryan et al., 2017).

Another potential limitation and validity threat of this research study is the curricular changes and improvements that occurred during the third year of undergraduate medical education at the research study site. Continuous quality improvement is a required component of the accreditation process for all United States medical schools (LCME, 2019). Faculty members and administrators are entrusted to engage in ongoing continuous quality improvement processes that establish short and long term programmatic goals. These planned goals result in the achievement of measurable outcomes used to improve programmatic quality and ensure effective monitoring of the medical education program's compliance with accreditation and other standards. The potential continuous quality improvements in curricular delivery at the research study site are considered a limitation to this research study design. The possible confounding quality improvement factors created variability in the curriculum and future assessment performance in the studied student records from one year to the next. There is a potential that

student performance changed from one year to the next due to these confounding factors and not the formative assessment program.

A further limitation of this research study was related to potential bias about the structure of the formative assessment program by the researcher who worked at the study site as an administrator at the time of the research study. Two steps were taken to combat potential biases. First, a thorough literature review was included in Chapter 2 that supported the rationale to carry out this research study. The literature review was used to build the structure supporting the importance of formative assessment programming in undergraduate medical education to help prepare undergraduate medical education students for their high-stakes summative examinations. This review of the literature showed the need for this research study, regardless of the researcher's opinions on the topic.

Another step taken to limit the researcher's bias was that an independent staff member at the study site provided de-identified data to the researcher. This process eliminated the potential for the researcher to have prior knowledge of any cases or include further data based on knowledge of individual students within the research study data set. Once the researcher officially requested the data set from the institution, the independent staff member, who has no affiliation with this research study, created a de-identified data set with the requested independent and dependent variables included. The independent staff member assigned random identification numbers to student names so that the researcher had no way to identify the individual students included in the data set. In another measure of bias control over the data set, the independent staff member destroyed the document linking identification numbers with student names and records. This measure eliminated the ability for later case exclusion based on prior knowledge of individual student records by the researcher. Once the de-identified data set

was delivered to the researcher from the independent staff member, the data set was unidentifiable and free from potentially biased case manipulation.

Internal and External Validity

Threats to both internal and external validity were inherent in this research study. According to Creswell (2015), threats to internal validity surround understanding the relationship between variables and determining if the existing relationship between the two variables was explainable by a linked causal relationship or if some other internal factor might be contributing to the relationship. A major internal validity threat in this research study was the potential lack of effort that students may have put forth while completing their formative assessments. The rationale for this research study was to investigate formative assessment programs at the undergraduate medical education level. There was the potential for students to discount the seriousness of the formative assessment program when completing the exams because they do not count towards a grade, which was an inherent threat to internal validity. Another threat to internal validity within this research study was the inability to capture each student's state-of-mind while completing the formative and summative assessments. Outside factors on test day can contribute to low performance for students, and there was no way to capture this information as it related to a student's final assessment scores analyzed in this research study.

External validity threats also existed within this research design. According to Creswell (2015), "threats to external validity are problems that threaten our ability to draw correct inferences from the sample data to other persons, settings, treatment variables, and measures. To overcome these threats, it is important to have strong research designs..." (p. 306). The studied data set encompassed a clearly defined set of inclusion criteria that included all performance records of third-year students at the research study site. No cases were excluded from this

research study. During a five-year window of time, all student records were included in the de-identified data set, but within those de-identified records are potential individual issues that impacted student performance that cannot be accounted for with the parameters of this research study. For example, some students experience the stress of summative assessments differently than others. This potential anxiety could have played a part in their summative assessment performance, but has little to do with the formative assessment program investigated in this research study. Other factors that contribute to external validity threats revolve around the potential differences in the physical setting from one formative assessment delivery to another and from one formative assessment to the summative assessment delivery from year to year. The researcher was unable to account for issues like room temperature, distractions in the testing room, the noise outside the testing room, or technical difficulties with some individual computers. These types of issues may have occurred on an individual basis and would not be measurable within the studied data set.

Ethical Issues

All requested and analyzed student information in this research study was part of regularly collected assessment data at the research study site, which exists as part of the overarching medical school assessment system. This research study did not ask students to provide information about their perception or satisfaction with the formative assessment program. Also, since the data were de-identified, the researcher did not know the identities of students and could not use this information in their capacity as an administrator at the medical school. The analysis of existing records was conducted after the independent staff member from the Office of Medical Education provided the de-identified data set for review of the formative assessment program at the medical school. From an ethical control standpoint, the researcher was

unable to edit the data set in any way based on relationships or knowledge obtained about individual student backgrounds, performance in other areas, or outcomes after graduation. The de-identified data set provided only the information to investigate the stated research questions without bias surrounding the data set inclusion or exclusion criteria. This issue is important to highlight because narrowing down the data set based on previous knowledge about individual students could potentially shift results to match the researcher's bias toward the outcome of the research study creating an ethical issue within the research study results.

There was potential for conflict of interest within this research study because the researcher was an administrative leader in the medical school from which the data was requested. The researcher stayed separate from the data-gathering process in every way adhering to the guidelines presented in the proposal and approved by the research study site and the University of New England institutional review boards. The independent staff member from the research study site separated the researcher from the data-gathering process through their de-identification of the data set. The only data used in this project's analysis was the data provided to the researcher at the time of the request, and no further cases were brought into the research study after the data set was produced. This action controlled for bias on the part of the administrator/researcher so that no confounding variables could be added into the analysis later in the research process. The practice at the medical school for a data request was to submit a written request to the Office of Medical Education for de-identified data. Standard operating procedures were in place at the medical school to allow researchers to separate themselves from their medical school roles to conduct ethical medical education research.

Conclusion and Summary

The purpose of this research study was to investigate the relationship between a comprehensive medical knowledge formative assessment program in undergraduate medical education and later summative clinical clerkship assessment performance. This research study utilized archival formative assessment program data to examine the impact and relationships formative assessments had on summative assessment performance. The independent variables included seven validated tools made available from the NBME to students and medical schools designed to inform students about their preparation for summative clinical clerkship subject examinations and attained medical knowledge. One primary goal of the researcher's approach to the research study included carrying out a reliable program evaluation with the ability to provide valid results to colleagues in undergraduate medical education. Of equal importance was ensuring a sound research design while limiting research bias and accounting for potential validity threats and limitations created by the researcher's role in the institution producing the data set. The scope of this research project investigated the performance of 332 third-year undergraduate medical students from the Northeastern Medical School. Student records were derived from cohort graduation dates ranging from the years 2016 to 2020.

The data analysis process focused on group differences related to formative and summative assessment performance. The analysis produced findings and tested hypotheses related to relationships between the independent and dependent variables. Undergraduate medical education students work to build medical knowledge before graduation and need consistent performance feedback related to their academic progress. Statistically significant findings surrounding the sustained utilization of a formative assessment program to aid in preparation and performance on high-stakes summative assessments and building of medical

knowledge would support the continued dedication of resources from an administrative leadership perspective, increase efficacy for the formative assessment program, and provide aid in medical knowledge development for future physicians with impending patient responsibilities in the near future.

CHAPTER 4

RESULTS

This quantitative program evaluation research study investigated undergraduate medical education student formative and summative assessment performance using archival data review. The purpose of the research study was to analyze the relationship of medical knowledge-based formative assessments to medical knowledge-based summative assessment performance. More specifically, the overarching research question for this research study is to what extent do medical knowledge-based formative assessments impact and relate to medical knowledge-based summative assessment performance in undergraduate medical education? The researcher focused the study on investigating if formative assessment programs have a statistically significant impact on summative assessment performance and the extent of the relationship between formative assessment performance and summative assessment performance in undergraduate medical education.

Analysis Method

This chapter provides an overview of the results of this research study. Below is a description of the data set and a descriptive analysis of the population. Once the de-identified data set was delivered to the researcher from the school's Office of Medical Education independent staff member in a Microsoft® Excel spreadsheet, the researcher coded the data set with variable titles. These titles included formative Family Medicine, formative Internal Medicine, formative Neurology, formative Obstetrics-Gynecology, formative Pediatrics, formative Psychiatry, and formative Surgery. The researcher next coded the summative assessment variables as follows: summative Family Medicine, summative Internal Medicine, summative Neurology, summative Obstetrics-Gynecology, summative Pediatrics, summative

Psychiatry, and summative Surgery. The final relevant variable code was the cohort the student belonged to pertinent to their anticipated graduation year. Northeastern Medical School increased enrollment progressively from 2016 through 2018. Earlier graduating classes were admitted with small class sizes, making the comparison of three cohorts of students with two cohorts of students valid because the clustered group sizes were relatively close to one another (177 students vs. 155 students).

Once the institutional review board applications from the University of New England and the research study site were approved, the data were organized, de-identified by the research study site independent staff member, and hand-delivered to the researcher with the needed password for access. The data set was saved onto a password-protected external hard drive using a Microsoft® Excel spreadsheet. The researcher accessed the spreadsheet using the provided password and began to investigate the provided data. The researcher next opened up IBM® SPSS Version 26 on their computer and transferred the data into IBM® SPSS Version 26 for further data analyses using the copy and paste function on their computer. SPSS is a statistical software program used to analyze data in the social science fields with descriptive, relational, and grouping variable analysis (Levesque, 2007). The researcher used IBM® SPSS because it is the preferred system of the researcher as they received their formal statistical training on it when learning to conduct statistical analysis. IBM® SPSS Version 26 is available for use at the researcher's institution for internal and external research purposes by faculty, staff, and students.

Two classifications were added to the data set when received by the researcher to conduct a further in-depth analysis of the data in IBM® SPSS Version 26. First, the researcher added a classification category that denoted the cohorts of students in the data set who were not exposed (Classes of 2016, 2017, & 2018) to the formative assessment program. These student records

were assigned a classification title of “non-formative assessment group.” The cohorts who were exposed (Classes of 2019 & 2020) to the formative assessment program were assigned a new variable classification title of “formative assessment group.” The creation of these new classifications occurred to make comparative analysis possible within the data set that would allow the researcher to investigate the efficacy of the formative assessment program. Secondly, the arithmetic mean was calculated by adding up all seven core clinical clerkship discipline formative assessment scores to create a new classification titled “average formative assessment performance.” Next, the researcher created a new classification for the arithmetic means (average score) for the seven core clinical clerkship discipline summative assessment scores entitled “average summative assessment performance.” The researcher created the arithmetic means classifications to investigate the overall performance of students across all formative and summative assessments collectively. The overall performance calculations allowed the researcher to analyze overall performance on the formative and summative assessments as a whole. These added calculations created a holistic view of formative and summative assessment program performance by students that accounted for variability in scores created by potential strengths and weaknesses, and/or interest and non-interest related to the individual clerkship disciplines.

Descriptive characteristics for each formative and summative assessment were calculated to investigate measures of central tendency, including mean and standard deviation. Next, the class year variable was recoded to describe each cohort. The researcher then conducted multiple statistical tests to analyze the data. The overarching research question for this research study is to what extent do medical knowledge-based formative assessments impact and relate to medical

knowledge-based summative assessment performance in undergraduate medical education? The following research sub-questions additionally guided this research study:

1. What effect do medical knowledge formative assessment programs have on individual clerkship assessment performance in undergraduate medical education?
2. What is the extent of the relationship between formative assessment performance and summative assessment performance in undergraduate medical education?

As part of this research study, multiple hypotheses were tested for further investigation of the research topic.

Hypotheses

H₀: Third-year medical students who completed formative assessments during their third year of undergraduate medical education score significantly higher on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education.

H₁: Third-year medical students who completed formative assessments during their third year of undergraduate medical education exhibit no difference in their performance on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education.

H₀: There is a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations.

H₁: There is no significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations.

The data analysis related to this research study resulted in the rejection of null hypothesis one. Through multiple independent samples *t*-tests, through data analysis the researcher found when students completed the formative assessment program during their third year of undergraduate medical education at the research study site, they scored significantly higher on their Family Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, and Surgery core clerkship summative subject examinations and their overall average summative subject examination performance than those students who did not complete formative assessments during their third year of undergraduate medical education. The only subject examination where completion of the formative assessment program did not show a significant increase in performance was in the Internal Medicine summative examination performance. Significant findings related to multiple independent samples *t*-tests supported hypothesis one in that exposure to the formative assessment program had a statistically significant impact on summative assessment performance. The data analysis related to this research study also allowed for the rejection of null hypothesis two. Through multiple Pearson's *r* correlation tests, the researcher found that third-year undergraduate medical students at the research study site displayed significant positive associations between formative and summative assessment performance in the Family Medicine, Internal Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, and Surgery clerkship disciplines, and overall average formative and summative subject examination performance. These significant findings related to hypothesis

two lead to the conclusion that formative assessment performance had a statistically significant relationship with summative assessment performance in this research study.

Presentation of Results

Research study variables included 332 student performance records. Variables in the data set included anticipated student graduation year upon entering the medical school, performance scores on the seven NBME clinical science mastery series formative assessments, and NBME summative clinical science subject examination performance, as presented below in Table 1. The research study utilized multiple independent samples *t*-tests to investigate differences between groups. The groups included three cohorts of third-year medical students not exposed to a formative assessment program (Classes of 2016, 2017, & 2018) and two cohorts exposed to a formative assessment program (Classes of 2019 & 2020). Investigating differences using independent samples *t*-tests provided the researcher with the ability to investigate the differences in performance on summative assessments of those students exposed to the formative assessment program and those not exposed to the formative assessment program. Analysis of the results allowed the researcher to extrapolate information on the efficacy of the formative assessment program.

The independent samples *t*-tests investigated the cohorts exposed and not exposed to the formative assessment programs in seven core clerkship disciplines, and an averaged assessment score. These clerkships included Family Medicine, Internal Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, and Surgery. The dependent variable encompassed student performance on the National Board of Medical Examiners summative subject examinations in these same seven core clerkship disciplines.

Table 1*Hypotheses One and Two Research Design*

	Hypothesis One			Hypothesis Two	
	Cohorts from Graduating Classes of 2016, 2017, 2018	Cohorts from Graduating Classes of 2019 & 2020	Investigating Differences between the two groups	Cohorts from Graduating Classes of 2019 & 2020	Investigating Relationships between the two variables
Clinical Clerkships (6 Week Blocks)	n = 177	n = 155	n = 332	n = 155	
Family Medicine			* Independent Variables (IV): Seven Formative Clinical Science Mastery Series Assessment Scores and average performance		
Internal Medicine	* No Formative Assessment Exposure	* Formative Assessment Exposure		* Formative Assessment Exposure	* Anticipated Statistical Analyses tests Pearson product-moment correlations. Investigating the significance and strength of relationships between performance on the formative (IV) and summative (DV) assessments and average performance
Neurology			* Dependent Variables (DV): Seven Summative Clinical Science Subject Assessment Scores and average performance		
Obstetrics-Gynecology					
Pediatrics	* Required to complete the Summative Assessments	* Required to complete the Summative Assessments	* Statistical Analyses: Independent Samples <i>t-tests</i> to investigate differences between groups exposed to IV and those not exposed to the IV before all completed DV	* Required to complete the Summative Assessments	
Psychiatry					
Surgery					

The research study also utilized Pearson's *r* correlation analyses. Benesty, Chen, Huang, and Cohen (2009) stated that the Pearson's *r* correlation provides research consumers with information about the relationship between two independent variables in the areas of statistical significance and strength. The Pearson's *r* correlations in this research study focused on measuring the significance and strength of the relationships between the formative NBME clinical science mastery series assessments and the summative NBME subject examinations.

Delivery of the assessment tools occurred at the research study location in the seven different core clerkship disciplines and for average performance across all formative and summative clerkship assessments, as outlined above in Table 1. Measuring the strength of the relationship between the formative and summative assessment performance scores had great importance in this research study. The correlation analysis in this research study provided the researcher with an understanding of the strength and significance of the relationships between the assessments in the same clerkship discipline and overall. The analysis also provided the researcher with information about the usefulness of the formative assessment tools in aiding summative assessment performance. Measuring the strength and significance of the relationship through Pearson's r correlations provides students, faculty, and administrators an understanding of where gaps in medical knowledge are and can direct preparation after completing their future summative assessments.

Descriptive Data

The research study population included assessment scores for 332 undergraduate third-year medical students spread across five different cohorts between the graduation year classes of 2016 to 2020, as described below in Table 2. The graduating Class of 2016 included 45 (13.6%) students, Class of 2017 included 60 (18.1%) students, Class of 2018 included 72 (21.7%) students, Class of 2019 included 79 (23.8%) students, and the Class of 2020 included 76 (22.9%) students. The formative assessment group included 177 (53.3%) of the students in the research study population, and the non-formative assessment group included 155 (46.7%) of the students in the research study population.

Table 2*Descriptive Data by Graduating Year*

Graduation Year	Number of Students	Formative Assessment
		Program Exposure
Class of 2016	45 (13.6%)	No Formative Assessment
Class of 2017	60 (18.1%)	Program Exposure
Class of 2018	72 (21.7%)	177 (53.3%)
Class of 2019	79 (23.8%)	Formative Assessment
Class of 2020	76 (22.9%)	Program Exposure
		155 (46.7%)
Total	332 (100%)	

The research study population, described below in Table 3, includes the formative assessment performance of 155 de-identified third-year medical students at the research study site from two cohorts of students (Classes of 2019 & 2020). Third-year undergraduate medical students at the research study institution completed the NBME clinical science mastery series formative assessments in seven core clerkships after their six-week clerkship blocks in the classes of 2019 and 2020. This instance of formative assessment use was the first delivery of the formative assessment program at the research study site in the third year of undergraduate medical education. Table 3 below provides detailed performance of both cohorts in the seven different clerkships' formative assessments and average student performance. The table provides data from the formative assessments across all clerkships and by cohort with mean performances and standard deviations in parentheses. The Class of 2019 NBME clinical science mastery series formative assessment average clerkship performance ranged from 59.70 (11.79) on the Surgery formative examination to 68.38 (13.45) on the Psychiatry formative examination. The average student performance across all seven formative examinations for the Class of 2019 was 64.95

(8.33). The Class of 2020 NBME clinical science mastery series formative assessment average performance ranged from 61.88 (11.88) on the Surgery formative examination to 70.04 (9.75) on the Psychiatry formative examination. The average student performance across all seven clerkship formative examinations for the Class of 2020 was 66.10 (7.61). When looking at performance on the NBME clinical science mastery series, formative assessment average performance by discipline, and across both the Classes of 2019 and 2020, the average performance ranged from 60.77 (11.85) on the Surgery formative examinations to 69.19 (11.78) on the Psychiatry formative examinations. The average student performance across all disciplines and both the Class of 2019 and 2020 was a 65.51 (7.98).

Table 3

Core Clerkship Formative Assessment Program Performance

Core Clerkship Discipline	Class of 2019 Mean (SD)	Class of 2020 Mean (SD)	Average Subject Mean (SD)
Family Medicine	65.62 (12.28)	66.54 (9.75)	66.07 (11.08)
Internal Medicine	62.51 (13.68)	63.88 (12.14)	63.18 (12.93)
Neurology	68.23 (13.51)	69 (14.52)	68.61 (13.97)
Obstetrics-Gynecology	66.10 (12.63)	67.17 (11.16)	66.63 (11.91)
Pediatrics	64.10 (14.59)	64.20 (10.83)	64.15 (12.84)
Psychiatry	68.38 (13.45)	70.04 (9.75)	69.19 (11.78)
Surgery	59.70 (11.79)	61.88 (11.88)	60.77 (11.85)
Student Average Score	64.95 (8.33)	66.10 (7.61)	65.51 (7.98)

Note SD represents the standard deviation

Third-year undergraduate medical students at the research study institution completed the NBME subject examinations as summative assessments in seven core clerkships at the end of

their third year of undergraduate medical education, but before entering their fourth year of medical school. Cohort performance from the seven different summative clerkship assessments is provided below in Table 4. The table displays data from the clerkship disciplines and cohort summative class average performances. This data is represented by showing both the mean performance and standard deviation in parentheses. The Class of 2016 summative NBME subject examination average performance ranged from a class average of 76.80 (9.66) on the Obstetric-Gynecology Clerkship examination to a class average of 85.11 (7.91) on the Psychiatry Clerkship examination. The Class of 2016's average score across all seven summative NBME subject examinations was 79.91 (7.52), which was the highest average performance across all five cohorts. The Class of 2017 summative NBME subject examination average performance ranged from a class average of 70.10 (8.93) on the Surgery Clerkship examination to a class average of 76.68 (8.26) on the Psychiatry Clerkship examination. The Class of 2017's average score across all seven summative NBME subject examinations was 74.34 (7.08), which was the lowest average performance across all five cohorts. The Class of 2018 summative NBME subject examination average performance ranged from a class average of 72.64 (7.76) on the Surgery Clerkship examination to a class average of 78.82 (7.16) on the Psychiatry Clerkship examination. The Class of 2018's average score across all seven summative NBME subject examinations was 76.56 (6.67), which was the second-lowest average performance across all five cohorts. The Class of 2019 summative NBME subject examination average performance ranged from a class average of 76.51 (7.69) on the Surgery Clerkship examination to a class average of 82.33 (6.78) on the Psychiatry Clerkship examination. The Class of 2019's average score across all seven summative NBME subject examinations was 79.07 (6.33), tied for second-highest average performance across all five cohorts. The Class of 2020's summative NBME subject

examination average performance ranged from a class average of 74.61 (6.95) on the Surgery Clerkship examination to a class average of 81.38 (6.15) on the Psychiatry Clerkship examination. The Class of 2020's average score across all seven summative NBME subject examinations was 79.07 (5.72), tied for second-highest average performance across all five cohorts. When focusing on the individual clerkship discipline, NBME subject examination average performances ranged from a class average of 74.67 (8.53) on the Surgery Clerkship examinations to a class average of 80.71 (7.61) on the Psychiatry Clerkship examinations. The average performance across all clerkships was 77.78 (6.83).

Table 4*Core Clerkship Summative Assessment Performance*

Core Clerkship Discipline	Class of 2016	Class of 2017	Class of 2018	Class of 2019	Class of 2020	Average Score
	Mean (SD)	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)
Family Medicine	79.02 (9.55)	74.37 (6.83)	78.35 (6.57)	80.91 (7.18)	80.59 (5.93)	78.84 (7.43)
Internal Medicine	80.62 (8.20)	74.17 (10.07)	76.60 (8.36)	78.57 (9.11)	77.50 (8.16)	77.38 (8.97)
Neurology	77.31 (7.55)	76.03 (8.42)	76.25 (7.54)	79.89 (6.65)	80.30 (6.53)	78.15 (7.48)
Obstetrics-Gynecology	76.80 (9.66)	75.33 (9.08)	76 (8)	77.27 (7.69)	79.33 (7.28)	77.05 (8.29)
Pediatrics	79.58 (9.17)	73.68 (7.83)	77.26 (9.33)	78.05 (7.68)	79.75 (6.98)	77.69 (8.38)
Psychiatry	85.11 (7.91)	76.68 (8.26)	78.82 (7.16)	82.33 (6.78)	81.38 (6.15)	80.71 (7.61)
Surgery	80.91 (8.65)	70.10 (8.93)	72.64 (7.76)	76.51 (7.69)	74.61 (6.95)	74.67 (8.53)
Student Average Score	79.91 (7.52)	74.34 (7.08)	76.56 (6.67)	79.07 (6.33)	79.07 (5.72)	77.78 (6.83)

Note SD represents the standard deviation

Hypothesis 1

Hypothesis one posited that third-year medical students who completed formative assessments during their third year of undergraduate medical education score significantly higher

on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education. Multiple independent samples *t*-tests were conducted to investigate the impact that the introduction of the formative assessment program had on summative assessment performance in each of the seven core clerkships and overall.

The researcher conducted eight individual independent samples *t*-tests. The *t*-tests allowed the researcher to investigate the impact exposure to formative assessments had on summative subject examination performance. Table 5 below provides information about third-year undergraduate medical student summative assessment performance separated by cohort exposure to the formative assessment program. Performance on the Family Medicine, Internal Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychology, and Surgery clerkships summative assessments, as well as overall average performance, are found below.

Table 5

Independent Samples t-tests: Efficacy of Formative Assessment Program

Clerkships	No Formative Assessment Program Exposure (n=177)	Formative Assessment Program Exposure (n=155)	<i>t</i> (330)	<i>P</i>	Null Hypothesis
	Mean(SD)	Mean(SD)			
Family Medicine	77.17 (7.74)	80.75 (6.58)	4.51	.001	Reject
Internal Medicine	76.80 (9.22)	78.05 (8.65)	1.27	<i>ns</i>	Accept
Neurology	76.45 (7.82)	80.09 (6.57)	4.56	.001	Reject
Obstetrics-Gynecology	75.98 (8.78)	78.27 (7.54)	2.54	.01	Reject
Pediatrics	76.64 (9.06)	78.88 (7.37)	2.46	.01	Reject
Psychology	79.69 (8.37)	81.86 (6.47)	2.61	.01	Reject
Surgery	73.88 (9.37)	75.57 (7.38)	1.81	.05	Reject
Student Average Score	76.66 (7.31)	79.07 (6.02)	3.26	.001	Reject

Note SD represents the standard deviation

After completing eight independent samples *t*-tests related to hypothesis one, the researcher rejected the null hypothesis after finding significant results in seven of the eight independent samples *t*-tests conducted. Exposure to the formative assessment program in this research study showed a statistically significant impact on the Family Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychology, and Surgery summative assessment performance, along with average student performance across all clerkship summative assessments. The Internal Medicine Clerkship was the only independent samples *t*-test that did not result in a significant difference in summative assessment performance when analyzing exposure and non-exposure to the formative assessment program.

The rejection of null hypothesis one is an essential finding for this research study. Seven of the eight individual independent samples *t*-tests displayed significant results in support of the independent variables. These findings indicate the formative assessment program had a significant impact on the summative assessment performance in seven of the eight independent samples *t*-tests conducted. These findings provide efficacy for the importance of the formative assessment program in this research study. These findings also provide support for investigating hypothesis two, which further examines the relationship between formative and summative assessment performance in this research study.

The researcher's findings support the efficacy of the formative assessment program. The results indicate that exposure to the formative assessment program did have a significant impact on summative assessment performance, based on the independent samples *t*-test findings. In six of the seven individual clerkship summative assessment performances, exposure to the formative assessment program was indicative of significantly higher scores on summative assessments, with the only clerkship not showing significant results being the Internal Medicine clerkship.

Most notable were the significant findings showing that overall average summative assessment performance, across all seven clerkships, was significantly higher in the groups exposed to the formative assessment program. Students may have individual strengths and weaknesses related to different clerkship disciplines. The significant differences in average performance across all clerkship disciplines concerning exposure to the formative assessment program show efficacy for the formative assessment program as a whole while controlling for individual strengths and weaknesses and/or interest and disinterest related to the different clerkship disciplines.

Hypothesis 2

Hypothesis two predicted that there was a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations. Results of Pearson's r correlations indicated that there was a significant positive association between formative and summative assessment performance in all core clerkship disciplines. All correlations were statistically significant at the .01 to .001 level. Table 6 below highlights the strength and significance level related to each clerkship disciplines Pearson's r correlation test. Data set 1 found in Appendix B illustrates the moderate, positive significant relationship between the Family Medicine Clerkship formative and summative assessment performance ($r(155) = .35, p < .001$) by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020. Data set 2 found in Appendix B illustrates the moderate, positive significant relationship between the Internal Medicine Clerkship formative and summative assessment performance ($r(155) = .41, p < .001$) by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020. Data set 3 found in Appendix B illustrates the moderate, positive significant relationship between the Neurology

Clerkship formative and summative assessment performance ($r(155) = .54, p < .001$) by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020. Data set 4 found in Appendix B illustrates the moderate, positive significant relationship between the OB-GYN Clerkship formative and summative assessment performance ($r(155) = .30, p < .01$) by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020. Data set 5 found in Appendix B illustrates the moderate, positive significant relationship between the Pediatrics Clerkship formative and summative assessment performance ($r(155) = .43, p < .001$) by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020. Data set 6 found in Appendix B illustrates the weak, positive significant relationship between the Psychiatry Clerkship formative and summative assessment performance ($r(155) = .22, p < .01$) by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020. Data set 7 found in Appendix B illustrates the moderate, positive significant relationship between the Surgery Clerkship formative and summative assessment performance ($r(155) = .44, p < .001$) by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020. Data set 8 found in Appendix B illustrates the strong, positive significant relationship between the overall average formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.

After completing eight Pearson's r correlation analyses tests related to hypothesis two, the researcher rejected null hypothesis two after finding significant results in all eight correlation analyses conducted. Performance on the formative assessments this research study showed statistically significant relationships with the Family Medicine, Internal Medicine, Neurology,

Obstetrics-Gynecology, Pediatrics, Psychology, and Surgery summative assessment performance, along with average student performance across all clerkships.

The rejection of null hypothesis two is another imperative finding for this research study. All eight individual Pearson's r correlation analyses tests displayed significant results. These findings indicate that formative assessment performance had a significant relationship with performance on summative assessments in the two cohorts of third-year undergraduate medical students at the research study institution. These findings provide further support for the importance of the formative assessment program in this research study. The strength of the relationships between the individual clerkship formative and summative assessments ranged from weak to strong. Average formative assessment performance and summative assessment performance displayed a strong relationship between the two variables. Table 6 below displays the interpretation of the strength of the relationships between formative and summative assessment performance (Akoglu, 2018).

Table 6

Strength of Relationship between Formative and Summative Assessment Performance

Core Clerkship Discipline	r - value	Strength of Relationship
Family Medicine	.35***	moderate
Internal Medicine	.41***	moderate
Neurology	.54***	moderate
Obstetrics-Gynecology	.30***	moderate
Pediatrics	.43***	moderate
Psychiatry	.22**	weak
Surgery	.44***	moderate
Student Average Score	.70***	strong

*** $p < .001$

** $p < .01$

Summary

The purpose of this research study was to investigate the impact and relationship formative assessment in third-year undergraduate medical education has on summative assessment performance. More specifically, the study investigated the impact of completing a formative assessment program throughout the academic year had on summative assessment performance in seven different core clerkships and overall average performance. Furthermore, the study investigated the relationship between formative assessment performance and summative assessment performance in the same seven core clerkships and overall average performance. Hypothesis one posited that medical students who completed formative assessments during their third year of undergraduate medical education would score significantly higher on their seven core clerkship subject examinations than those students who did not complete formative assessments during their third year of undergraduate medical education. The second hypothesis predicted a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations.

After statistical testing, the researcher found statistically significant results allowing for the rejection of both null hypotheses. Exposure to the formative assessment program had a statistically significant impact on summative assessment performance. Furthermore, formative assessment performance had a statistically significant relationship with summative assessment performance. Chapter Five provides a short introduction, the interpretation of the findings, implications that the findings from this research study indicate, recommendations for future activities related to this research, recommendations for further study related to this research, and the conclusion.

CHAPTER 5

CONCLUSION

The researcher's role as an administrator in undergraduate medical education was a major contributing factor to the origins of this research study. An equally important contributing factor was a literature investigation into the importance of formative assessment. The research literature on this topic yielded extensive information about foundations for the importance of formative assessment, but very little specific information related to formative assessment programs in undergraduate medical education. Most notable in the literature search was the research from Black and Wiliam (2017), who concluded that formative assessment is essential in providing the greatest likelihood for the highest student achievement. The current study builds upon the work of Black and Wiliam in the area of undergraduate medical education formative assessment with the analysis of the impact and relationship formative assessment programming has on summative assessment performance.

The purpose of this study was to investigate the impact and relationship formative assessment in third-year undergraduate medical education has on summative assessment performance. The overarching research question for this research is to what extent do medical knowledge-based formative assessments impact and relate to medical knowledge-based summative assessment performance in undergraduate medical education? The following research sub-questions additionally guide this research study:

1. What effect do medical knowledge formative assessment programs have on individual clerkship assessment performance in undergraduate medical education?
2. What is the extent of the relationship between formative assessment performance and summative assessment performance in undergraduate medical education?

During their third year of medical school, as part of their regular undergraduate medical education assessment program, undergraduate medical education students at the research study site in the classes of 2019 and 2020 completed clerkship discipline-specific medical knowledge formative subject examination assessments after completing six weeks of their clinical clerkship experience. These clerkships included Family Medicine, Internal Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, and Surgery. After the third year of medical school, but before starting the fourth year of undergraduate medical education, the same students completed clerkship discipline-specific medical knowledge summative subject examinations, which played a role in determining their clerkship grades. Hypothesis one posited that medical students who completed formative assessments during their third year of undergraduate medical education score significantly higher on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education. The researcher conducted multiple independent samples *t*-tests to investigate the impact that the introduction of the formative assessment program had on summative assessment performance in the seven core clerkships and overall. After completing eight independent samples *t*-tests related to hypothesis one, the researcher rejected the null hypothesis after finding significant results in seven of the eight independent samples *t*-tests conducted. Exposure to the formative assessment program in this research study showed a statistically significant impact on the Family Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychology, and Surgery summative assessment performance, along with average student performance across all clerkship summative assessments. The Internal Medicine Clerkship was the only discipline in which the independent samples *t*-test did not result in a significant difference in summative assessment performance when analyzing exposure and non-

exposure to the formative assessment program. Hypothesis two predicted that there was a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations. Results of Pearson's r correlations indicated that there was a significant positive association between formative and summative assessment performance in all core clerkship disciplines. All correlations were statistically significant at the .01 to .001 level.

The testing of both hypotheses resulted in significant findings. Results showed that formative assessment programming significantly impacted summative assessment performance and that formative assessment performance had significant relationships with summative assessment performance at the research study site. When shared widely, the significant findings related to this research study could be helpful for multiple stakeholders in undergraduate medical education. These stakeholders include medical education administrators when making decisions about budget allocations within their organizations related to student assessment, undergraduate medical education faculty, specifically clerkship faculty, to include a similar formative assessment program within their curricular structure to aid in student performance and attainment of medical knowledge, and medical students in order for them to embrace a formative assessment program offered by their undergraduate medical education institution designed to help enhance their summative assessment performance and knowledge-base for future clinical practice.

Interpretation of Findings

The overarching research question for this study was to what extent do medical knowledge-based formative assessments impact and relate to medical knowledge-based summative assessment performance in the third year of undergraduate medical education?

Hypothesis one predicted that medical students who completed formative assessments during their third year of undergraduate medical education score significantly higher on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education. Hypothesis one was informed by the previous work of Chang and Wimmers (2016), who found that regular formative assessment aids in the monitoring process of student achievement. Hypothesis two predicted that there is a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clinical science subject examinations. Hypothesis two was informed by the work of Konopasek, Norcini, and Krupat (2016) who found that formative assessment in medical education cannot be a series of single events, but a natural process that students embrace over time or in a series of events.

Hypothesis One Interpretation

Hypothesis one posited that medical students who completed formative assessments during their third year of undergraduate medical education score significantly higher on their seven core clerkship subject examinations than those students who do not complete formative assessments during their third year of undergraduate medical education. Statistical analysis using independent samples *t*-tests resulted in significant findings for hypothesis one, which showed that medical students who completed formative assessments during their third year of undergraduate medical education scored significantly higher on the Family Medicine, Neurology, OB-GYN, Pediatrics, Psychiatry, and Surgery Clerkship subject examinations, and in average subject examination performance than those students who did not complete formative assessments during their third year of undergraduate medical education. Summative assessment performance was significantly higher for the cohorts who completed the formative assessment

program as opposed to those who did not complete the formative assessment program specifically the results of the eight independent samples *t*-tests displayed significant findings in six of the seven clerkship disciplines, and overall average performance of students who were exposed to the formative assessment program scored significantly higher on their summative assessments than students who were not exposed to the formative assessment program.

When analyzing the impact formative assessment programming had on summative assessment performance, only one clerkship (Internal Medicine) did not show significant results. There was no significant impact on Internal Medicine summative assessment performance based on those third-year undergraduate medical students not exposed and exposed to the Internal Medicine formative assessment. There is a potential confounding variable or outside causal factor not related to the variables analyzed in this research study. Instead, the confounding factor potentially impacting the non-significant finding for the Internal Medicine Clerkship formative and summative assessment is inherent to the nature of the Internal Medicine clerkship itself. The educational objectives and assessment material found on the Internal Medicine formative and summative assessment are also taught and assessed within the six other third-year clerkships (Gao, Askew, Violato, Manthey, Burns, & Vallevand, 2019). Internal Medicine is the foundation for the other six core clerkships at the research study site. The continued reinforcement of medical knowledge in the other disciplines could have led to an increase in general medical knowledge over time that would be most relevant to the Internal Medicine formative and summative assessments. Based on the work of Gao, Askew, Violato, Manthey, Burns, and Vallevand (2019), if material from the Internal Medicine Clerkship is taught during the other six clerkships at the research study site, the analysis of the relationship between formative and summative assessment could be skewed by those factors and contributing to the only non-

significant finding in the research study, surrounding the Internal Medicine Clerkship formative and summative assessments.

The findings related to hypothesis one and the efficacy of the formative assessment program at the research study site are concordant with literature presented earlier in this research study surrounding the importance of formative assessment. Specifically, Moore (2018) identified the need for learning through formative assessment for students as they work toward their career goals while building a robust medical knowledge-base. The significant findings related to hypothesis one align with Moore's (2018) findings, which support the significant impact formative assessment programming had on summative assessment performance and the attainment of medical knowledge available to the third-year undergraduate medical students as they progress into their medical career after graduating. The formative assessment program provides students with needed information about their current medical knowledge-base related to the particular clerkship discipline they are completing the formative assessment for during the academic year. Students at the study site then take the feedback they received after completing their formative assessment and use that information to target areas of weakness to improve upon before completion of the summative subject examination in the same clerkship disciplines at the end of the academic year. Because hypothesis one was rejected showing students exposed to the formative assessment program performed better on all but one of their clerkship summative assessments and overall average performance and because the formative and summative assessments are medical knowledge-based questions, the improved performance supports the attainment of further medical knowledge when exposed to the formative assessment program which is needed as students' progress into medical professionals with patient care responsibility.

The current study also aligns with Deiglmayr's (2018) research which discussed the importance of the immediate impact formative assessment has on the undergraduate medical student in the development of their medical knowledge strengths and weaknesses, along with findings from Dolin, Black, Harlen, and Tiberghien (2018) who posited that planned formative assessment systems ensure students are receiving quality information about their current knowledge-base and where gaps in that knowledge-base are identifiable. The formative assessment program analyzed at the research study site provided students with immediate performance feedback after their formative assessments were completed. This feedback allowed students to assess their strengths and weaknesses before progressing to their summative assessments, where students exposed to the formative assessments at the study site performed significantly higher than those students not exposed to the formative assessment program.

Hypothesis Two Interpretation

To investigate hypothesis two, the researcher performed multiple Pearson's r correlation tests focusing on the relationship between formative and summative assessments within discipline-specific clerkships taken by third-year medical students at the study site. Specifically, the researcher predicted there was a significant relationship between third-year undergraduate medical education student performance on their seven core clerkship formative assessments and their summative clerkship assessments. This prediction was informed by the work of Chang and Wimmers (2016), who posited that regular formative assessment aids in the monitoring process of student achievement. The hypothesis was also informed by the work of Konopasek, Norcini, and Krupat (2016) who concluded that formative assessment in medical education could not be a series of single events, but a natural process that students embrace over time or in a series of events. Both works support the structure of the formative assessment program at the study site

and the predication of significant relationships between formative and summative assessment performance in the same discipline-specific clerkships and overall. Results indicated that there was a significant positive association between formative and summative assessment performance in all clerkship disciplines, including Family Medicine, Internal Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, Surgery, and overall average performance. The strength of relationships of the Pearson's r correlations ranged from weak to strong ($r = .22$ to $.70$). Findings related to hypothesis two align with the previous work of Brenner, Bird, and Willey (2017), who posited that standardized curricular formative assessments provided the ability to predict later summative assessment performance. The authors focused on singular formative assessments related to licensure examinations. In contrast, the current study focused on the formative assessment program as a whole. Still, both studies displayed results about the vital relationship formative assessments can have on summative assessment performance.

The statistical correlation analysis performed in the research study supports the efficacy of the formative assessment program, as demonstrated by the statistically significant findings surrounding hypothesis two. All Pearson's r correlation tests resulted in positive significant correlations, which shows significant support for the formative assessment program. The findings associated with hypothesis two show a significant relationship between undergraduate medical student performance on formative and summative assessments. This indicates that third-year undergraduate medical students at the research study site had a significant relationship in their performance on formative assessments and summative assessments in the same clerkship disciplines and overall performance. When analyzing the performance of the 155 third-year undergraduate medical students who completed the formative assessment program at the research study site, there were positive statistically significant relationships between each

clerkship's formative and summative assessment performance, with relationship strength ranging from weak to strong. The strongest significant relationship analyzed was overall average formative and summative assessment performance ($r(155) = .70, p < .001$). This finding is indicative of the strength of the relationship between formative assessment programming in supporting overall summative assessment performance. These findings provided strong support for the formative assessment program holistically because they encompass average student performance across Family Medicine, Internal Medicine, Neurology, OB-GYN, Pediatrics, Psychiatry, and Surgery clerkship formative and summative assessments. Although individual clerkship correlation strengths varied from weak to moderate as outlined above in Table 6, the overall average performance correlation analysis displayed the strongest correlation between two variables. This finding strongly supports the importance of utilizing the overall formative assessment program in undergraduate medical education because the strongest correlation was related to the relationship between the overall average formative assessment and overall average summative assessment performance.

The findings supporting the formative assessment program's significant relationships with summative assessments align with the prediction of the researcher, first hypothesized during the study design phase of the research study. The significant findings in the current research study are also supported by previous work from Houston and Thompson (2017), who posited that assessment events and structure are the key components to improving summative assessment performance. The formative assessment program used at the research study site is built into the curriculum encountered by all third-year medical students who complete the formative assessments at the conclusion of their six-week clerkship blocks. As described in Houston and Thompson (2017), the scheduled formative assessment events and structure of the formative

assessment program throughout the third-year of the study site's curriculum, culminating in summative assessments at the end of the academic year, allowed administrators, faculty, and students to plan and embrace the formative assessment program as a regular part of their schedules while working towards improved summative assessment performance after encountering the formative assessment program. The findings also align with previous work from Batool, Asim, Shah, and Chughtai (2018), who presented findings that reflected the important positive influence formative assessment has on later summative assessment performance in medical education. The significant findings related to hypotheses one and two supported the impact and relationship that the formative assessment program had on later summative assessment performance.

The results of this research study provide direct answers to the research questions posed by the researcher during the initial phase of study design. The findings related to hypothesis one showed the significant impact the formative assessment program had on summative assessment performance, with significant findings the Family Medicine, Neurology, OB-GYN, Pediatrics, Psychiatry, and Surgery Clerkship subject examinations, and in average subject examination performance than those students who did not complete formative assessments during their third year of undergraduate medical education which showed the effect medical knowledge formative assessment programs have on individual clerkship assessment performance in undergraduate medical education. The findings related to hypothesis two which showed positive significant relationships between Family Medicine, Internal Medicine, Neurology, OB-GYN, Pediatrics, Psychiatry, and Surgery Clerkship subject examinations, and in the average subject examination, formative and summative examination performance displayed the extent of the relationship

between formative assessment performance and summative assessment performance in undergraduate medical education.

Student Perception

The current research study findings extend and support previous research focused on the perception of students, faculty, and administrators concerning the importance of formative assessment. Notably, Langendyk (2006) discussed the difficulty struggling students presented when trying to assess their performance after completing assessments. The findings in the current research study provide support for student participation in a formative assessment program to gain a better understanding of their related knowledge, strengths, and weaknesses going into summative assessments and for later patient care duties. This is supported by the positive significant Pearson's r correlations between formative and summative assessment performance in the Family Medicine, Internal Medicine, Neurology, OB-GYN, Pediatrics, Psychiatry, and Surgery Clerkships and overall average performance at the research study site. By utilizing the feedback formative assessment program provided student participants, they gain the needed information about their current knowledge in a specific discipline. The results of the current research study also offer support for Kibble's (2007) research, which promoted the frequent usage of formative assessments as a way to ensure more reliable performance on later summative assessments. This support is most evident from the findings in hypothesis two, which showed multiple highly significant positive correlations between formative and summative assessment performance.

Faculty Perception

From a faculty perspective, Close (2017), reported that faculty perception of formative assessment was an important factor determining faculty support during the delivery of feedback

to students. Faculty play an important role in the support and inclusion of all aspects of medical education curricula (Thomas, Kern, Hughes, & Chen, 2016). The current research study findings support the efficacy and relationship between formative assessment programming and summative assessment performance based on the findings related to hypotheses one and two previously discussed. These significant findings could bolster faculty support to incorporate formative assessment programming within their institutions with the hope of finding similar results to the current research study.

Administrative Leadership Perception

The current research study findings show third-year undergraduate medical students who participated in formative assessment programming performed better on summative assessments than students not exposed to formative assessment programming lend strong support to previous research surrounding the need for institutional and administrative support of formative assessment programming. Most notably, Hersh and Keeling (2013) concluded that a firm commitment from the institution is needed surrounding assessment to promote the education and learning process of students. The research study findings also align with the work of Palmer and Devitt (2018), who posited that formative assessment as part of curricular design showed a strong positive impact on student performance and engagement in the educational process. Bolman and Deal (as cited in Gallos, 2006) note the importance that organizational leaders play in supporting and sustaining change within their organization. Support from leaders must be a consistent effort that includes financial, staffing, and space support for a robust formative assessment system in this change initiative. Without support from leadership, the planned action will not succeed in attaining the import goals related to the proposed program because of the financial, staff allocation, and physical space needed to carry out the program detailed in the

current research study. This is supported by Hersh and Keelings' (2013) discussion about the need for strong institutional support for the strength of the educational process. The significant finding related to Hypothesis one and two surrounding efficacy of formative assessment programming on summative assessment performance and the relationship between formative and summative assessment performance in the current research study displayed the significant impact formative assessment programming has on summative assessment performance in undergraduate medical education students. These findings should gain the support of other institutional leaders for formative assessment. Formative assessment programming requires further support from undergraduate medical education institutional leaders related to financial and staff resources. These leaders must consider the significant positive impact formative assessment programs can have on medical knowledge attainment and career advancement, which is important for future physicians.

Implications

Undergraduate medical education students, faculty, and administrative leadership may use the significant results related to the efficacy of formative assessment programs and relationship between formative and summative assessment performance found in this study in their decision-making process related to the usage, development, and support of formative assessment programming within their institutions. Students may use the results of this study when making decisions about the efforts they plan to dedicate to completing their required formative assessments and the amount of trust they place in using the formative assessment results for adapting their study plans for summative assessments because of the positive significant correlations found in this study. Faculty members may use the results of this study to dedicate educational sessions to formative assessments and including dedicated formative

assessment time when planning longitudinal curricula. Administrative leadership may use the results of this research study when making decisions about the dedication of financial and staff resources surrounding the utilization of formative assessment tools and delivery. Because of the significant findings at the research institution that displayed support for the efficacy of formative assessment programming in hypothesis one and the positive significant relationship between formative and summative assessment performance supporting hypothesis two, the research study has the potential to start a much broader conversation in the undergraduate medical education community about how best to structure formative assessments at their schools. The significant findings from this research need to be communicated broadly to the medical education community. From there, medical education leaders can make informed decisions about choosing the appropriate types of formative assessment to include in their third-year curriculum. The significant findings in this research study provide support for other institutions to investigate the appropriateness of a similar formative assessment program at their institution to help improve performance on medical knowledge-based assessments and further develop medical knowledge in their students and future graduates. As noted previously, the undergraduate medical education accrediting body requires formative assessment planning as part of medical school curricula but does not prescribe how it is incorporated (LCME, 2019). The current research findings support further cultivation of formative assessment programs in the undergraduate medical education field because the significant results showed that formative assessment program exposure could significantly improve summative assessment performance and medical knowledge. The research study findings that displayed support for the efficacy of formative assessment programming in hypothesis one and the positive significant relationship between formative and summative assessment performance supporting hypothesis two may lead to more defined roles of formative

assessment programming in undergraduate medical education because of the support these findings displayed for the inclusion of such a program. The next section discusses recommendations for further action related to formative assessment programming in undergraduate medical education.

Recommendations for Action

The significant findings displayed in this research study provide support for the efficacy of formative assessment programming in hypothesis one, and the positive significant relationship between formative and summative assessment performance supporting hypothesis two in this research study has great importance for recommending further action from stakeholders. Most notably, the research study findings provide support for the implementation of formative assessment programs in undergraduate medical education institutions. Because formative assessment programming showed a significant positive impact and relationship in summative assessment performance, formative assessment program development deserves further investigation and support at undergraduate medical education institutions. Following the foundational support for the importance of formative assessment in the literature, undergraduate medical institutions should utilize the findings from this research study in coordination with the literature and be able to extrapolate a plan of action related to developing a formative assessment program at their institution. Formal formative assessment programming allows institutions to provide their students with the best opportunity to succeed when completing their required formative assessments, later licensure examinations, and building their knowledge base in the medical field. The author recommends incorporating formative assessment programming where it fits best for the institution. The programming should be molded into the needs of the institution to provide further support for students.

Recommendations for Further Study

McConnell, Harms, and Saperson (2016) found that student formative assessment must inform knowledge-base and aid students in achieving their goals during the education process. In further research focused on student outcome goals, Menéndez, Napa, Moreira, and Zambrano (2019) stressed the importance of the formative assessment process. Their work focused on the important role that formative assessment has in relation to the education process while also emphasizing how a continuous formative evaluation of student knowledge will improve eventual knowledge goals. Most recently, Tekian, Harden, Cook, Steinert, Hunt, and Norcini (2020) concluded that, with the import role formative assessment plays in medical education and attainment of student knowledge goals, a shift is needed to create more frequent formative assessment opportunities. The findings presented in the previously described studies provide foundational support for the current research study findings which focused on the importance of formative assessment and culminated in significant findings supporting the use of a structured ongoing formative assessment system occurring throughout the undergraduate medical education academic year resulting in improved summative assessment performance at the end of the academic year. The current research study findings further align with the work from the studies mentioned above in this section with support for the significant impact and relationship formative assessment has on summative assessment performance and knowledge acquisition through independent samples *t*-tests and correlation analysis focusing on the formative and summative assessment tools at the study site. The research study findings support the important role the creation and utilization of formative assessment programs have in undergraduate medical education by showing third-year undergraduate medical student exposure to formative

assessment programming results in significantly increased summative assessment performance in multiple clerkship disciplines and overall average assessment performance. Furthermore, the current research study displayed findings that performance on these impactful formative assessment tools show positive significant relationships with summative assessment tools in the same clerkship disciplines and overall.

The researcher's findings in this research study support the previously presented theoretical framework with the implementation and use of formative assessment programming in undergraduate medical education. The framework focused on adult learning theory and andragogy first presented by Knowles (1968), who proposed that adult learners are more independent and search for motivation internally related to knowledge acquisition goals. These same learners base much of their knowledge acquisition on what they believe they need to know to be successful (Knowles, 1989). Medical education as a whole revolves around self-directed learning because of the vast amount of information a physician must acquire throughout their career. This situation leaves medical students with inherent knowledge gaps when building their medical knowledge-base. Medical knowledge acquisition is an important component of undergraduate medical education. Medical students utilize attained medical knowledge to complete their required assessments and progress through their medical education curriculum. The formative assessment program, analyzed in this research study, allowed the adult learners (medical students) to assess their knowledge-base and motivate themselves to fill knowledge gaps based on their formative assessment outcomes before completing summative assessments which were supported by seven significant independent samples *t*-tests which showed students exposed to a formative assessment program performed better on summative assessments than students not exposed to the formative assessments as detailed in Table 5 above. The findings in

this research study also provided support for medical student participation in formative assessment programming by leading students to medical knowledge they need to be successful. The tenants of andragogy and adult learning theory application apply to the research study findings because the formative assessment program allows medical students to arrive at their conclusions about what they need to know through participation in the formative assessment program. Students then work to fill knowledge gaps independently throughout their clinical clerkship education and leading up to the summative assessment.

The research study findings also support the second pillar of the previously mentioned theoretical framework, which refers to components of transformational learning theory developed by Mezirow (1978). Mezirow discussed the importance of student reflection on their educational experiences to validate and reformulate their understanding of obtained knowledge and learning. Based on the current research study's findings, formative assessment programming has a significant impact and relationship with summative assessment performance. These findings, specifically detailed in Table 6 above, show the strength of the positive significant relationships between all formative and summative assessment performance within the Family Medicine, Internal Medicine, Neurology, OB-GYN, Pediatrics, Psychiatry, and Surgery Clerkships and overall average performance. Undergraduate medical students can reformulate learning plans to achieve learning goals after receiving feedback through formative assessment, aligning with transformational learning theory and the second pillar of the presented theoretical framework for this research study. Students can critically assess if they are meeting the desired levels of competency related to their learning goals and objectives of clerkships in advance of summative assessments. These two theories provided the foundation and lens for the current research study, stressing the importance of structured formative assessment programs with

ongoing assessment of one's ability while reformulating plans for more significant knowledge gains.

As previously described, the research study occurred at one medical school with a limited number of available student records, which made generalizability and controls for external validity threats within the research study a potential limitation. Another possible limitation was the number of cases analyzed in the population for this research study. The researcher investigated all available cases from the study site medical school, which came into existence in the last decade. This was completed by analyzing how cohorts of third-year undergraduate medical students at the study site location performed on discipline-specific clerkship summative assessments and overall average performance across summative assessments. Independent samples *t*-tests were used to investigate differences among cohorts who were and who were not exposed to a formative assessment program. Increasing the population size within this study could also increase the internal validity of this research study. A larger population size could enhance the study by adding the ability for the researcher to conduct predictive analyses between multiple independent and dependent variables, which are only valid to perform with a much larger set of cases than were available for this research study. The researcher recommends the study be replicated at a legacy school with larger class sizes to investigate formative assessment program performance on a larger scale.

The LCME (2019) requires that all undergraduate medical schools undergo continuous quality improvement practices. This research study produced significant findings surrounding formative and summative assessment that allows for multiple areas of potential future research. This research study's findings show that medical students who completed formative assessments during their third year of undergraduate medical education scored significantly higher on the

Family Medicine, Neurology, OB-GYN, Pediatrics, Psychiatry, and Surgery Clerkship subject examinations, and in average subject examination performance than those students who did not complete formative assessments during their third year of undergraduate medical education.

Also, the research found that multiple Pearson's r correlation tests indicated there was a significant positive association between formative and summative assessment performance in all conducted correlation testing. These results applied to all clerkship disciplines, including Family Medicine, Internal Medicine, Neurology, Obstetrics-Gynecology, Pediatrics, Psychiatry, Surgery, and overall average performance. The strength of the Pearson's r correlations ranged from weak to strong, but all findings were significant.

The researcher recommends that future researchers design studies that encompass different populations focusing on formative assessment programs in similar and different areas of medical education. This research study focused on one medical school's third-year assessment system. There are multiple other components of undergraduate medical education assessment that could benefit from a formative assessment program. Most notably are areas where formative assessment programming could have a positive impact on clinical skill development before graduation while working with patients.

Clinical skills development education is an essential learning tool for those seeking careers in patient care (Veesart & Johnson, 2020). Both formative and summative assessments are required within the undergraduate medical education clinical setting (LCME, 2019). The development of a formalized formative assessment program focusing on clinical assessment, in a similar way the current research study focused on the attainment of medical knowledge, could be beneficial for undergraduate medical education curriculum development and further the skill level of undergraduate medical education students.

Another potential future research project related to this research study could compare the formative assessment system described here with the same system at another undergraduate medical school. The other school would need to have different admission criteria to investigate the role incoming metrics have on later performance on formative and summative assessments. Such a research study could investigate differences in the overall population and their impact on efficacy and relationship of formative assessment programming and summative assessment outcomes. Another interesting project design could utilize survey research to investigate how other undergraduate medical schools conduct their required formative assessment operations and their perceptions of the efficacy of their program. Another interesting research project would be to investigate the relationship between formative assessment programs in medical education and other professional schools with required summative assessments as part of their field of study. This type of study would allow researchers to investigate the role of formative assessment programs outside of undergraduate medical education.

Finally, researchers should expand upon the current work of this research study at the research study site institution to investigate future relationship strength between formative and summative assessment performance as higher numbers of students join the population. Further analysis of the same variables used in this research study is relevant to investigate to find greater validity in the current research study's significant findings. If, as population size grows, the significant results from this study stay at current significance levels or continue to strengthen, those findings would further support the current study's outcomes that formative assessment programs significantly impact and relate to summative assessment performance in undergraduate medical education. If the significance levels begin to decrease or dissipate, this will show the current study's findings are only valid when analyzing a specific sample of students in

undergraduate medical education. As greater numbers of student performance records are added to the data analysis, the inclusion of more descriptive variables should be added to the de-identified data set. One imperative area to focus on is diverse populations of medical students. This further investigation could focus on formative assessment programs and summative performance outcomes within different groups. Identification of underrepresented groups related to culture, testing accommodation, admission status, prior work experience, and previous education could be important to investigate. This analysis could provide support for more targeted formative assessment programs for different groups within medical education.

A final potential next step to the current research project would be to develop a formative assessment perception survey that includes both qualitative and quantitative questions. The survey delivery may consist of medical students before their first encounter with the formative assessment programming and again after completing their summative assessments to look for themes and perceptions related to the formative assessment program, needed improvements, and positive outcomes from the student perspective.

Conclusion

The goal of this research was to investigate the impact and relationship formative assessment programs had on summative assessment performance and the attainment of medical knowledge. The greater focus of the researcher within this project and throughout their doctoral program education was to build strong transformational leadership abilities. The entire doctoral program and dissertation process is a true transformation of self. Bass (2008) concluded that transformational leaders gain active engagement of those they lead through their commitment to a goal, esteem within an organization, and charisma when delivering communication about shared goals. The transformation of self through doctoral study has strengthened the researcher's

commitment to the goals of the institution through the analysis of the formative assessment program.

From a transformational leadership perspective, communicating the implications of this research study and their importance to the Northeastern Medical School is imperative. The charismatic leader can empower followers with needed information while also increasing their support, regardless of any inherent conflict or resistance to the critical findings associated with this research project (Marion & Gonzales, 2014). The researcher intends to present findings related to this research study to organizational leaders, faculty, and students to ensure continued support for the formative assessment program. The communication of the information must embody the importance the researcher placed on the completion of the research project to gain the buy-in from the entire school community. The transformational leader strives to achieve organizational success through enacting change related to their goals. The findings related to this research must inform the student population and medical school in future decision-making processes.

The research study findings are imperative for continued administrative support from a financial and resource perspective related to the formative assessment program. Equally important is the buy-in from future students who will encounter the formative assessment program in their third year of medical school at the research study site. Based on the current research study findings, students, faculty, and medical school administration leadership can feel confident that exposure to the formative assessment program has a positive impact on their summative assessment performance. Going through formative assessment programming allows students to gain an understanding of where they are in their medical knowledge acquisition before they complete their summative assessments. The formative assessment programming also

allows student identification of gaps they may need to fill before moving into the next phase of their medical career.

The final component of transformational leadership relates to the transformation of the greater community, which, in the case of this research study, is the medical education community (Bass, 2008). The researcher hopes to present findings from this research study to the medical education community for further reflection and discussion. Insights gained through this research study will provide other medical school leaders with quantitative data regarding the importance of implementing formative assessment programs and the relationship formative assessments can have on summative assessment performance.

According to the LCME (2019), accreditation standards require the delivery of formative assessment in undergraduate medical education, but the accreditors do not prescribe how the assessments are utilized in their definition of formative assessment and the degree of needed inclusion. The findings from this research study can direct how formative assessment aligns with undergraduate medical education operations. The findings from this research study also allow other institutions to gain insight into what prescribed methods of formative assessment are available and most helpful to their students. This research is a starting point for investigating formative assessment programs in medical education. Using the knowledge discovered and presented in this research study, the medical education community and research study site will have a better understanding of formative assessment programming and the impact it has on medical student summative assessment performance and medical knowledge acquisition.

References

- AAMC: Association of American Medical Colleges (n.d.). Retrieved from <http://www.aamc.org/medicalschoools.htm>
- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine, 18*(3), 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anfara, V. A., & Mertz, N. T. (2015). *Theoretical frameworks in qualitative research*. (2nd Ed). Thousand Oaks, Calif: SAGE Publications.
- Al Kadri, H. M., Al-Moamary, M. S., Magzoub, M. E., Roberts, C., & van der Vleuten, C. (2011). Students' perceptions of the impact of assessment on approaches to learning: a comparison between two medical schools with similar Curricula. *International Journal of Medical Education, 2*, 44–52.
- Batool, H., Asim, U., Shah, S., & Chughtai, A. (2018). Formative assessment as a performance predictor for summative assessment in undergraduate medical students: A comparative correlational analysis in the subject of pathology. *Biomedica, 34*(3), 205.
- Bass, B. M. (2008). *The Bass handbook of leadership: Theory, research, and managerial applications*. New York: Free Press.
- Beaudoin, M. (2012). Institutional leadership: Transformative change or disruptive technology. In: Moore, M. (ed.) *Handbook of Distance Education* (3rd edition). Mahwah, NJ: Lawrence Erlbaum Associates.

- Benesty J., Chen J., Huang Y., & Cohen I. (2009). Pearson Correlation Coefficient. In: *Noise Reduction in Speech Processing*. Springer Topics in Signal Processing, vol 2. Springer, Berlin, Heidelberg
- Benjamin, L., Benjamin, G., Benjamin, C., & Selfridge, N. (2018). STEPS in the basic clinical skills curriculum, *Medical Teacher*, 40(6), 569-572.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Black, P., & Wiliam, D. (2017). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551-575.
- Bloom, Benjamin S. (1968). *Learning for mastery*. Los Angeles, USA: University of California Press.
- Bok, H. G., Teunissen, P. W., Favier, R. P., Rietbroek, N. J., Theyse, L. F., Brommer, H..., & Jaarsma, D. A. (2013). Programmatic assessment of competency-based workplace learning: when theory meets practice. *BMC Medical Education*, 13, 123-132.
- Bolman, L. & Deal, T. (2006). Reframing Change: Training, realigning, negotiating, grieving, and moving on. In J.V. Gallos (Ed.), *Organization development* (pp. 447-469). San Francisco, CA: Jossey-Bass.
- Brenner, J.M., Bird, J.B., & Willey, J.M. (2017). Formative assessment in an integrated curriculum: Identifying at-risk students for poor performance on usmle step 1 using nbme custom exam questions. *Academic Medicine*, 92, S21-S25.
- Buelin, J., Ernst, J. V., Clark, A. C., D.T.E., Kelly, D. P., & DeLuca, V. W. (2019). Formative evaluation techniques. *Technology and Engineering Teacher*, 78(5), 21-23. Retrieved from

<http://ezproxy.rowan.edu/login?url=https://search.proquest.com/docview/2173643041?accountid=13605>

- Burke, K. (2010). *Balanced assessment*. Bloomington, Ind.: Solution Tree Press.
- Chang, E.K., & Wimmers, P.F. (2016). Effect of repeated/spaced formative assessments on medical school final exam performance. *Health Professions Education*, 3(1), 32 - 37.
- Chappuis, J. (2014). Thoughtful assessment with the learner in mind. *Educational Leadership*, 71(6), 20-26.
- Close, B. (2017). *Faculty perceptions of formative assessment and implementation practices in preclinical medical education: A Q method study*. Retrieved from ProQuest LLC: ERIC Number: ED579533.
- Creswell, J.W. (2015). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (5th ed.). Boston, MA: Pearson.
- Deiglmayr, A. (2018). Instructional scaffolds for learning from formative peer assessment: effects of core task, peer feedback, and dialogue. *European Journal of Psychological Education*, 33(1), 185.
- Deeley, S. J. (2018). Using technology to facilitate effective assessment for learning and feedback in higher education. *Assessment & Evaluation in Higher Education*, 43(3), 439-448.
- Demir, M. (2018). Using online peer assessment in an instructional technology and material design course through social media. *Higher Education*, 75(3), 399–414.
- Dolin J., Black P., Harlen W., & Tiberghien A. (2018). Exploring relations between formative and summative assessment. In: Dolin J., Evans R. (eds) *Transforming Assessment. Contributions from Science Education Research*, vol 4. Springer, Cham

- Doukas, D. & Volpe, R. (2018). Why pull the arrow when you cannot see the target? framing professionalism goals in medical education. *Academic Medicine*, 93(11), 1610-1612.
- Downing, S.M. & Yudkowsky, R. (2009). Introduction to Assessment in the Health Professions. In Downing, S.M. & Yudkowsky, R. (Eds.). *Assessment in Health Professions Education* (pp. 1 -20). NY, NY: Routledge.
- Eltorai, A. E. M. (2013). Publishing medical schools' usmle step 1 scores: increase preclinical education accountability and national standards. *Advances in Medical Education and Practice*, 4, 135-136. doi: <http://dx.doi.org.une.idm.oclc.org/10.2147/AMEP.S51646>
- Elmahdi, I., Al-Hattami, A., & Fawzi, H. (2018). Using technology for formative assessment to improve students' learning. *Turkish Online Journal of Educational Technology*, 17(2), 182-188.
- Farrell, L., Bourgeois-Law, G., Buydens, S., & Regehr, G. (2019). Your goals, my goals, our goals: The complexity of co-constructing goals with learners in medical education. *Teaching and Learning in Medicine*, 31(4), 370-377.
DOI:10.1080/10401334.2019.1576526
- Ferris, H. & O'Flynn, D. (2015). Assessment in medical education; what are we trying to achieve? *International Journal of Higher Education*, 4(2), 139-144.
- Gao, H., Askew, K., Violato, C., Manthey, D., Burns, C., & Vallevand, A. (2019). Does clerkship rotation sequence affect performance on national board of medical examiners (nbme) clinical subject examinations and united states medical licensing examination (usmle) step 2 clinical knowledge (ck) examination?. *Medical Science Educator*, 29(3), 763-770.

- Gardner-Gletty, D.C. (2002). *Student Learning Outcomes Assessment and Faculty Work in Higher Education: A Comparative Study of Assessment Implementation in a Research Extensive Setting*. (Doctoral Dissertation). Retrieved from University of Wisconsin Madison Library- Dissertation Holdings:
https://elpa.education.wisc.edu/docs/WebDispenser/elpa-dissertation-abstracts/02summer_gardner-gletty.doc?sfvrsn=3
- Gonsalves, C. & Zaidi, Z. (2016). "Hands in medicine: understanding the impact of competency-based education on the formation of medical students' identities in the United States." *Journal of Educational Evaluation for Health Professions*, 13(31), 1-5.
- Gruppan, L., Ten Cate, O., Lingard, L. A., Teunissen, P. W. & Kogan, J. R. (2018). Enhanced requirements for assessment in a competency-based, time-variable medical education system. *Academic Medicine*, 93(3), S17-S21.
- Gullo, C.A., McCarthy, M.J., Shapiro, J.I., et al. (2015). Predicting medical student success on licensure exams. *Medical Science Educator*, 25(4), 447-453.
- Guskey, T. R. (2005). *Formative Classroom Assessment and Benjamin S. Bloom: Theory, Research, and Implications* presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada, April 11-15, 2005.
- Harrison, C., Könings, K., Dannefer, E., Schuwirth, L., Wass, V., van der Vleuten, C. (2016). Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. *Perspectives on Medical Education*, 5(5), 276-284.
- Hayek, S., Lane, S., Fluck, M., Hunsinger, M., Blansfield, J., & Shabahang, M. (2017). Ten year projections for us residency positions: will there be enough positions to accommodate the

- growing number of u.s. medical school graduates. *Journal of Surgical Education*, 75(3), 546- 551.
- Hersh, R. H., & Keeling, R. P. (2013). Changing Institutional Culture to Promote Assessment of Higher Learning. *Learning Institute for Learning Outcomes Assessment, Occasional Paper # 17*, (pp 1-16). Retrieved from www.learningoutcomesassessment.org
- Hill, W. S. (2014). *School professionals' perceptions of formative assessment as related to student achievement* (Order No. 3682157). Available from Education Database; ProQuest Central. (1657424664). Retrieved from <http://ezproxy.rowan.edu/login?url=https://search.proquest.com/docview/1657424664?accountid=13605>
- Houston, D., & Thompson, J. N. (2017). Blending formative and summative assessment in a capstone subject: 'It's not your tools, it's how you use them.' *Journal of University Teaching & Learning Practice*, 14(3), 2.
- Kayaalp, M. (2018). Modes of De-identification. AMIA ... Annual Symposium proceedings. *AMIA Symposium*, 2017, 1044–1050.
- Khalil, M.K., Hawkins, H.G., Crespo, L.M., & Buggy, J. (2018). The design and development of prediction models for maximizing students' academic achievement. *Medical Science Educator*, 28(1), 111-117.
- Kibble, J. (2007). Use of unsupervised online quizzes as formative assessment in a medical physiology course: effects of incentives on student participation and performance. *Advances in Physiology Education*, 31(3), 253-60.
- Knowles, M. S. (1968). Andragogy, not pedagogy. *Adult Leadership*, 16(10), 350–352.

- Knowles, M. (1989). *The making of an adult educator: An autobiographical journey* (Ed.). San Francisco, CA: Jossey-Bass.
- Konopasek, L., Norcini, J., & Krupat, E. (2016). Focusing on the formative: Building an assessment system aimed at student growth and development. *Academic Medicine*, 91(11), 1492-1497.
- Kotter, J.P. (2012). *Leading Change*. Boston, MA: Harvard Business Review Press.
- Kumar, A. D., Shah, M. K., Maley, J. H., Evron, J., Gyftopoulos, A., & Miller, C. (2015). Preparing to take the USMLE Step 1: A survey on medical students' self-reported study habits. *Postgraduate Medical Journal*, 91(1075), 257.
doi:<http://dx.doi.org.une.idm.oclc.org/10.1136/postgradmedj-2014-133081>
- Langendyk, V. (2006). Not knowing that they do not know: self-assessment accuracy of third-year medical students. *Medical Education*, 40, 173-179
- Lee, M.W., Johnson, T.R. & Kibble, J. (2017). Development of statistical models to predict medical student performance on the usmle step 1 as a catalyst for deployment of student services. *Medical Science Educator*, 27(4), 663.
- Leedy, P.D. & Ormrod, J. E. (2010). *Practical Research: Planning and Design*, Ninth Edition. NYC: Merrill.
- Leggio, L.E. & Albritton, T. A. (2015). Feedback. In Pangaro, L. N. & McGaghie, W.C. (Eds.) *Handbook on Medical Student Evaluation and Assessment*. Syracuse, NY: Gegensatz Press.
- Levesque, R. (2007). *SPSS Programming and Data Management: A Guide for SPSS and SAS Users (4th ed.)*. Chicago, Illinois: SPSS Inc. ISBN 978-1-56827-390-7.
- Liaison Committee on Medical Education (2019). Retrieved from <http://lcme.org/about>

- Marion, R. & Gonzales, L. D. (2014). *Leadership in education: organizational theory for the practitioner* (2nd Ed). Long Grove, Ill: Waveland Press.
- McConnell, M. M., Harms, S., & Saperson, K. (2016). Meaningful feedback in medical education: Challenging the "failure to fail" using narrative methodology. *Academic Psychiatry: The Journal of the American Association of Directors of Psychiatric Residency Training and the Association for Academic Psychiatry*, 40(2), 377-379. DOI:10.1007/s40596-015-0370-5
- Menéndez, I. Y. C., Napa, M. A. C., Moreira, M. L. M., & Zambrano, G. G. V. (2019). The importance of formative assessment in the learning teaching process. *International Journal of Social Sciences and Humanities*, 3(2), 238-249.
- Merriam, S. B. (2001). Andragogy and self-directed learning: pillars of adult learning theory. *New Directions for Adult & Continuing Education*, 2001(89), 3.
- Mezirow, J. (1978). "Perspective transformation." *Adult Education Quarterly*, 28 (2), 100–110.
- Miller, S. T. (2009). *Formative computer-based assessments: The potentials and pitfalls of two formative computer-based assessments used in professional learning programs* (Order No. NR48227). Available from Education Database; ProQuest Central. (305048958). Retrieved from <http://ezproxy.rowan.edu/login?url=https://search.proquest.com/docview/305048958?accountid=13605>
- Mills, M.R. & Bettis, P. J. (2015). Using Multiple Theoretical Frameworks to Study Organizational Change and Identity. In Anfara Jr., V. A. & Mertz, N.T. (Eds.), *Theoretical Frameworks in Qualitative Research* (2nd ed) (pp. 96 – 118). Los Angeles, CA: SAGE.

- Minor, S., Stumbar, SE., & Bonnin, R. (2019). Impact of the introduction of the clinical science mastery series on family medicine nbme subject examination performance. *Family Medicine*, 51(8), 687-690. <https://doi.org/10.22454/FamMed.2019.253577>
- Mohr, L. (1990). *Understanding significance testing*. Thousand Oaks: SAGE Publications.
- Moore, D. E. (2018). Assessment of learning and program evaluation in health professions education programs. *New Directions for Adults and Continuing Education*, 157, 51 -64.
- Morrison, C., Smith, L., Ross, L., et al. (2016). Learning through self-assessment: Investigating the relationship between performance on the nbme® clinical science mastery series self-assessments and clinical science subject examinations. *Medical Science Educator*, 26, 665–672. <https://doi-org.ezproxy.rowan.edu/10.1007/s40670-016-0330-z>
- Mulliner, E. & Tucker, M. (2017). Feedback on feedback practice: perceptions of students and academics. *Assessment & Evaluation in Higher Education*, 42(2), 266-288.
- National Board of Medical Examiners ® (2019). *Clinical Science Mastery Series*. Retrieved from <https://www.nbme.org/>
- National Board of Medical Examiners. (2017). NBME guide to the Subject Examination Program <https://www.nbme.org/pdf/SubjectExams/SubExamInfoGuide.pdf>. Accessed April 17, 2019.
- Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students*. Boston, MA: Pearson/Allyn & Bacon.
- Norcini, J., Anderson, M. B., Bollela, V., Burch, V., Costa, M. J., Duvuvier, R., Hays, Mackay, M.F.P., Roberts, T., & Swanson, D. (2018). 2018 consensus framework for good assessment. *Medical Teacher*, 9, 1-8.

- O'Leary M, Scully D, Karakolidis A, & Pitsia V. (2018). The state-of-the-art in digital technology-based assessment. *European Journal of Education*, 53, 160–175.
- Palmer, E., & Devitt, P. (2014). The assessment of a structured online formative assessment program: A randomized controlled trial. *BMC Medical Education*, 14, 8 - 12.
- Palmen, L. N., Vorstenbosch, M. A. T. M., Tanck, E., & Kooloos, J. G. M. (2015). What is more effective: a daily or a weekly formative test? *Perspectives on Medical Education*, 4(2), 73–78.
- Pangaro, L. N. (2015). A Primer of Evaluation Terminology: Definitions and Important Distinctions in Evaluation. In Pangaro, L.N. and McGaghie, W. C. (Eds.). *Handbook on Medical Student Evaluation and Assessment* (pp. 13 -26). North Syracuse, NY: Gegensatz Press.
- Popham, W. J. (2013). *Transformative Assessment*. Boston, MA: Houghton Mifflin Harcourt.
- Raupach, T., & Schuelper, N. (2018). Reconsidering the role of assessments in undergraduate medical education. *Medical Education*, 52(5), 464-466. DOI:10.1111/medu.13543
- Ravitch, S. & Riggan, M. (2017). *Reason & rigor: How conceptual frameworks guide research*. 2nd edition. Thousand Oaks, CA. SAGE Publications.
- Rushton, A. (2005). Formative assessment: a key to deep learning? *Medical Teacher*, 27 (6), 509-513.
- Ryan, M., Bishop, S., Browning, J., Anand, R., Waterhouse, E., Rigby, F., ... Colbert-Getz, J. (2017). Are scores from nbme subject examinations valid measures of knowledge acquired during clinical clerkships? *Academic Medicine : Journal of the Association of American Medical Colleges*, 92(6), 847–852.
<https://doi.org/10.1097/ACM.0000000000001535>


- Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Saunders, R. (2014). Improving student performance through enhanced feedback. *International Journal of Management Education*, 12(2), 167-176.
- Sawilowsky, S. S., & Hillman, S. B. (1992). Power of the independent samples *t*-test under a prevalent psychometric measure distribution. *Journal of Consulting and Clinical Psychology*, 60(2), 240-243. doi: <http://dx.doi.org/10.1037/0022-006X.60.2.240>
- Scriven, M. (1967). The methodology of evaluation. R.W. Tyler, R M. Gagne, M. Scriven (eds.), In *Perspectives of curriculum evaluation*, (pp.39-83). Chicago, IL: Rand McNally.
- Sestoft, P. (2011). Organizing research data. *Acta Veterinaria Scandinavica*, 53 (1) , S2.
DOI:10.1186/1751-0147-53-S1-S2
- Shetach, A. (2012). Conflict leadership: Navigating toward effective and efficient team outcomes. *The Journal for Quality and Participation*, 35(2), 25-30.
- Shute, V.J. & Kim, Y.J. (2013). Formative and Stealth Assessment. In J.M. Spector et al. (eds.), *Handbook of Research on Educational Communications and Technology*, Springer Science Business Media, NY.
- Tekian, A., Harden, R. M., Cook, D.A., Steinert, Y., Hunt, D., & Norcini, J. (2020). Managing the tension: From innovation to application in health professions education. *Medical Teacher*, 42(3), 333-339.
- Thomas, P. A., Kern, D. E., Hughes, M. T., & Chen, B. Y. (Eds.). (2016). Curriculum development for medical education: a six-step approach. JHU Press.

- United States Medical Licensing Examination ® (2019). *What is USMLE?*. Retrieved from <https://www.usmle.org/>
- Veesart, A., Johnson, K. (2020). Using the Two-Minute Assessment With Debriefing to Optimize Traditional Patient Care Clinical Experiences, *Nursing Education Perspectives*, 41(2), 132-133.
- Watling, C. (2016). The uneasy alliance of assessment and feedback. *Perspectives on Medical Education*, 5(5), 262-264.
- Weaver-Hightower, M. B. (2014, February 11). *Conceptual frameworks* (part 1) [video file]. Retrieved from <https://www.youtube.com/watch?v=guUdGZWgKdw>
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-9.
- Yan, X., Su, X., & World Scientific (Firm). (2009). *Linear regression analysis: Theory and computing*. Singapore: World Scientific Pub. Co.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *The International Journal of Higher Education Research*, 45(4), 477-501.

Appendix A

1/9/2020

<https://eirb.rowan.edu/eirb/Doc/0/D0I2C4N7V9CKT10759AKD319CF/#fromString.html>



** This is an auto-generated email. Please do not reply to this email message.
The originating e-mail account is not monitored.
If you have questions, please contact your local IRB office **

DHHS Federal Wide Assurance Identifier: FWA00007111
IRB Chair Person: Ane Johnson
IRB Director: Sreekant Murthy
Effective Date: 11/5/2019

eIRB Notice of Approval

STUDY PROFILE

Study ID: Pro2019000795
Title: Understanding the Importance of Formative Assessment Programs in Undergraduate Medical Education

Principal Investigator:	Matthew Gentile	Study Coordinator:	Matthew Gentile
Co-Investigator(s):	There are no items to display	Other Study Staff:	There are no items to display
Sponsor:	There are no items to display	Approval Cycle:	Twelve Months
Risk Determination:	Minimal Risk	Device Determination:	Not Applicable

Review Type: Exempt **Expedited Category:** There are no items to display **Exempt Category:** 1
4(ii)

Subjects: **Specimens:** **Records:**

CURRENT SUBMISSION STATUS

Submission Type:	Research Protocol/Study	Submission Status:	Approved
Approval Date:	11/5/2019	Expiration Date:	

Continuation Review Required:

Pregnancy Code:	Not Applicable	Pediatric Code:	Not Applicable	Prisoner Code:	Not Applicable
------------------------	----------------	------------------------	----------------	-----------------------	----------------

Protocol:	Gentile Protocol EIRB.docx	Consent:	There are no items to display	Recruitment Materials:	There are no items to display
------------------	----------------------------	-----------------	-------------------------------	-------------------------------	-------------------------------

* **Retrospective Chart Review.** If applicable, records may be accessed to review information dating: **From:** **To:**

* **Study Performance Sites:**

CMSRU Campus 401 Broadway Camden, NJ 08103

<https://eirb.rowan.edu/eirb/Doc/0/D0I2C4N7V9CKT10759AKD319CF/#fromString.html>

1/2

1/9/2020

<https://eirb.rowan.edu/eirb/Doc/0/D0I2C4N7V9CKT10759AKD319CF/fromString.html>**ALL APPROVED INVESTIGATOR(S) MUST COMPLY WITH THE FOLLOWING:**

1. Conduct the research in accordance with the protocol, applicable laws and regulations, and the principles of research ethics as set forth in the Belmont Report.
- 2a. **Continuing Review:** Approval is valid until the protocol expiration date shown above. To avoid lapses in approval, submit a continuation application at least eight weeks before the study expiration date.
- 2b. **Progress Report:** Approval is valid until the protocol expiration date shown above. To avoid lapses, an annual progress report is required at least 21 days prior to the expiration date.
3. **Expiration of IRB Approval:** If IRB approval expires, effective the date of expiration and until the continuing review approval is issued: **All research activities must stop unless the IRB finds that it is in the best interest of individual subjects to continue. (This determination shall be based on a separate written request from the PI to the IRB.) No new subjects may be enrolled and no samples/charts/surveys may be collected, reviewed, and/or analyzed.**
4. **Amendments/Modifications/Revisions:** If you wish to change any aspect of this study after the approval date mentioned in this letter, including but not limited to, study procedures, consent form(s), investigators, advertisements, the protocol document, investigator drug brochure, or accrual goals, you are required to obtain IRB review and approval prior to implementation of these changes unless necessary to eliminate apparent immediate hazards to subjects. This policy is also applicable to progress reports.
5. **Unanticipated Problems:** Unanticipated problems involving risk to subjects or others must be reported to the IRB Office (45 CFR 46, 21 CFR 312, 812) as required, in the appropriate time as specified in the attachment online at: <http://www.rowan.edu/som/hsp/>
6. **Protocol Deviations and Violations:** Deviations from/violations of the approved study protocol must be reported to the IRB Office (45 CFR 46, 21 CFR 312, 812) as required, in the appropriate time as specified in the attachment online at: <http://www.rowan.edu/som/hsp/>
7. **Consent/Assent:** The IRB has reviewed and approved the consent and/or assent process, waiver and/or alteration described in this protocol as required by 45 CFR 46 and 21 CFR 50, 56, (if FDA regulated research). Only the versions of the documents included in the approved process may be used to document informed consent and/or assent of study subjects; each subject must receive a copy of the approved form(s); and a copy of each signed form must be filed in a secure place in the subject's medical/patient/research record.
8. **Completion of Study:** Notify the IRB when your study has been completed or stopped for any reason. Neither study closure by the sponsor nor the investigator removes the obligation for submission of timely continuing review application, progress report or final report.
9. The Investigator(s) did not participate in the review, discussion, or vote of this protocol.
10. **Letter Comments:** *There are no additional comments.*

CONFIDENTIALITY NOTICE: This email communication may contain private, confidential, or legally privileged information intended for the sole use of the designated and/or duly authorized recipients(s). If you are not the intended recipient or have received this email in error, please notify the sender immediately by email and permanently delete all copies of this email including all attachments without reading them. If you are the intended recipient, secure the contents in a manner that conforms to all applicable state and/or federal requirements related to privacy and confidentiality of such information.

<https://eirb.rowan.edu/eirb/Doc/0/D0I2C4N7V9CKT10759AKD319CF/fromString.html>

2/2



Institutional Review Board
Mary DeSilva, Chair

Biddeford Campus
11 Hills Beach Road
Biddeford, ME 04005
(207)602-2244 T
(207)602-5905 F

Portland Campus
716 Stevens Avenue
Portland, ME 04103

To: Matthew Gentile
Cc: Bill Boozang, Ed.D.
From: Lliam Harrison, M.A., J.D. CIM
Date: February 14, 2020
Project # & Title: 20.02.06-001 Understanding the Importance of Formative Assessment Programs in Undergraduate Medical Education

The Institutional Review Board (IRB) for the Protection of Human Subjects has reviewed the materials submitted in connection with the above captioned project and has determined that the proposed work is exempt from IRB review and oversight as defined by 45 CFR 46.104 (d)(4).

Additional IRB review and approval is not required for this protocol as submitted. If you wish to change your protocol at any time, including after any subsequent review by any other IRB, you must first submit the changes for review.

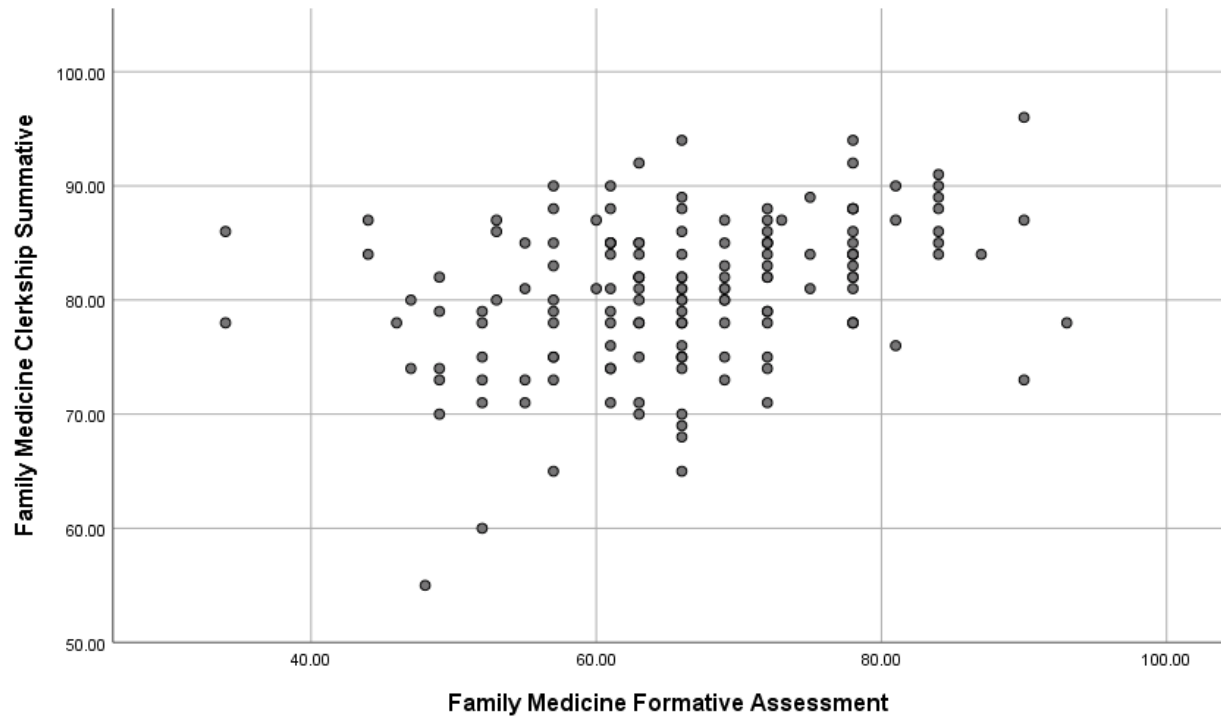
Please contact Lliam Harrison at (207) 602-2244 or wharrison@une.edu with any questions.

Sincerely,

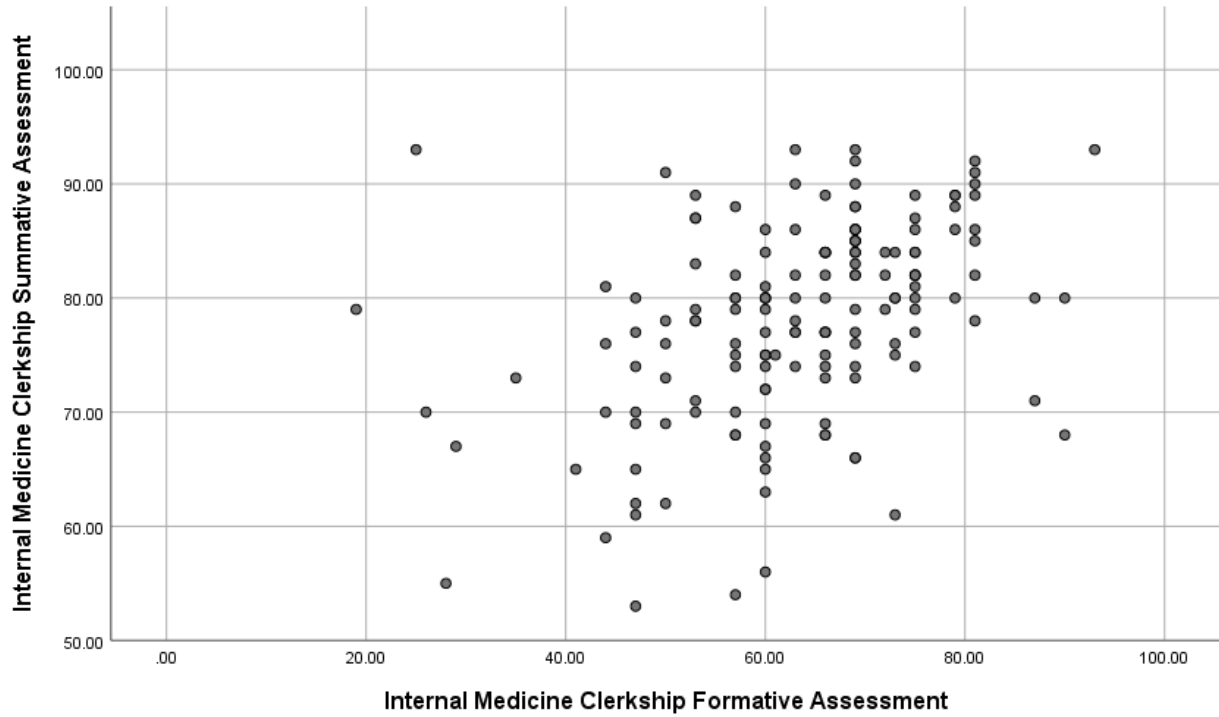
William R. Harrison, M.A., J.D. CIM
Director of Research Integrity

IRB#: 20.02.06-001
Submission Date: 01/13/20
Status: Exempt, 45 CFR 46.104 (d)(4)
Status Date: 2/14/20

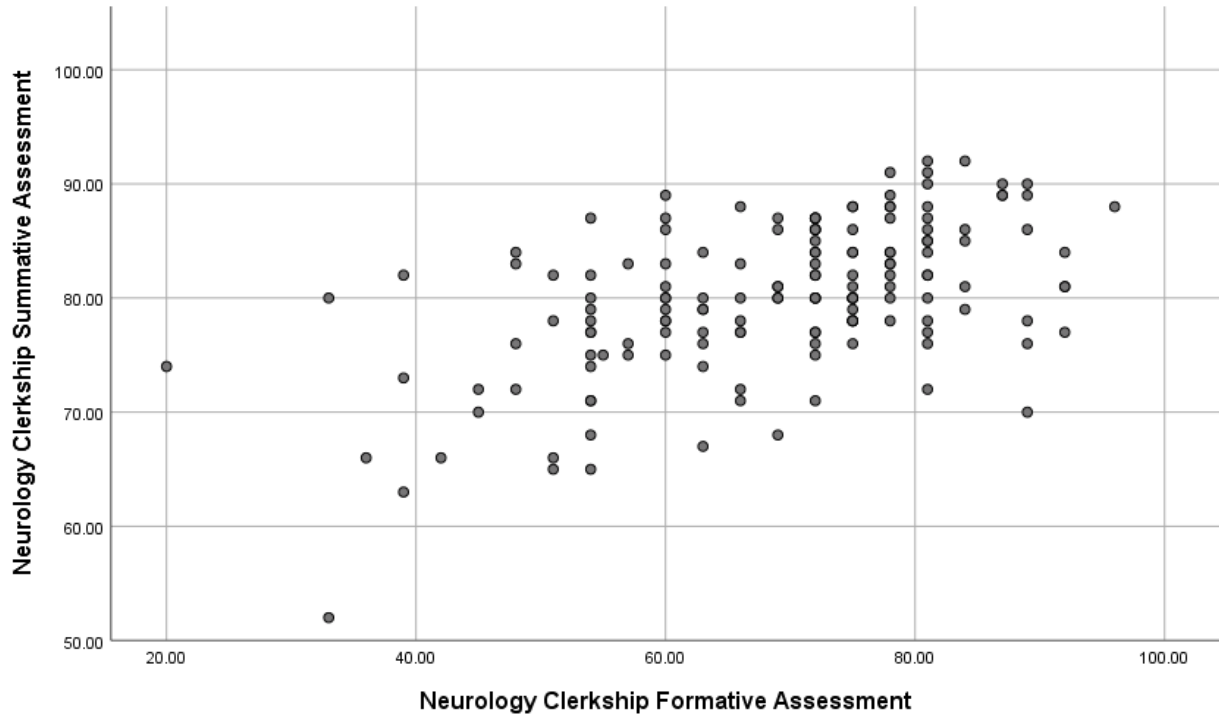
Appendix B



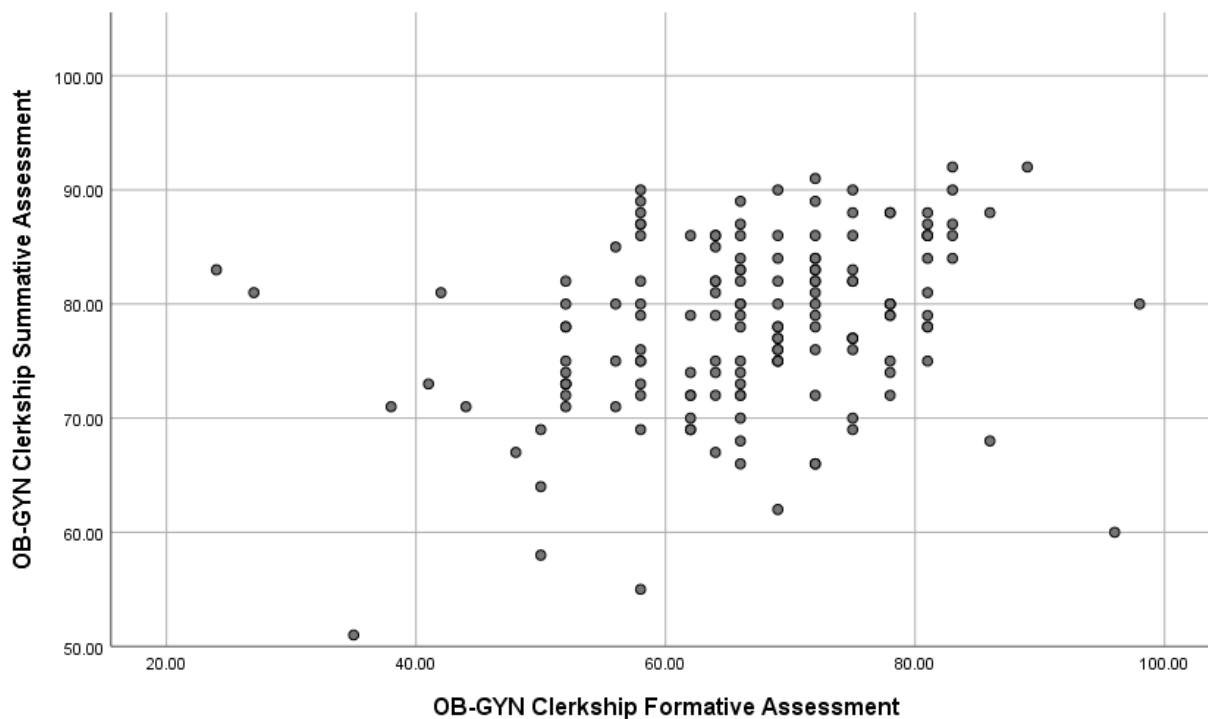
Data Set 1. Family Medicine Clerkship scatter plot. This data set illustrates the relationship between the Family Medicine Clerkship formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.



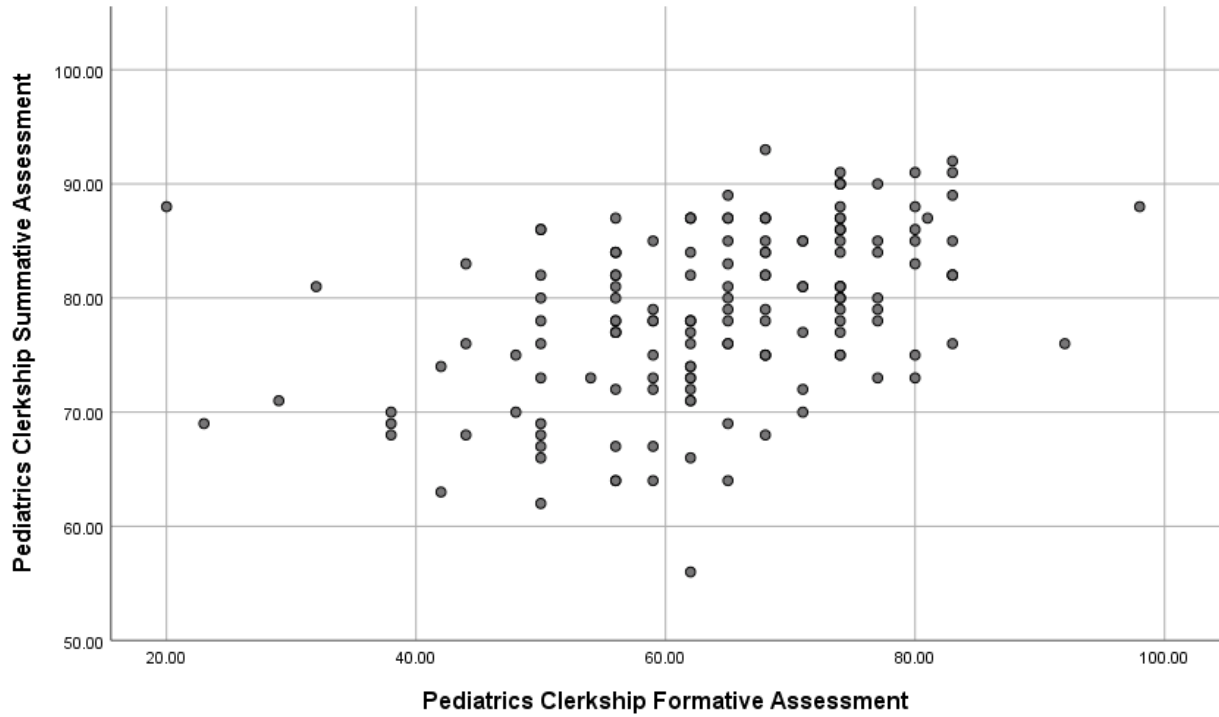
Data Set 2. Internal Medicine Clerkship scatter plot. This data set illustrates the relationship between the Internal Medicine Clerkship formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.



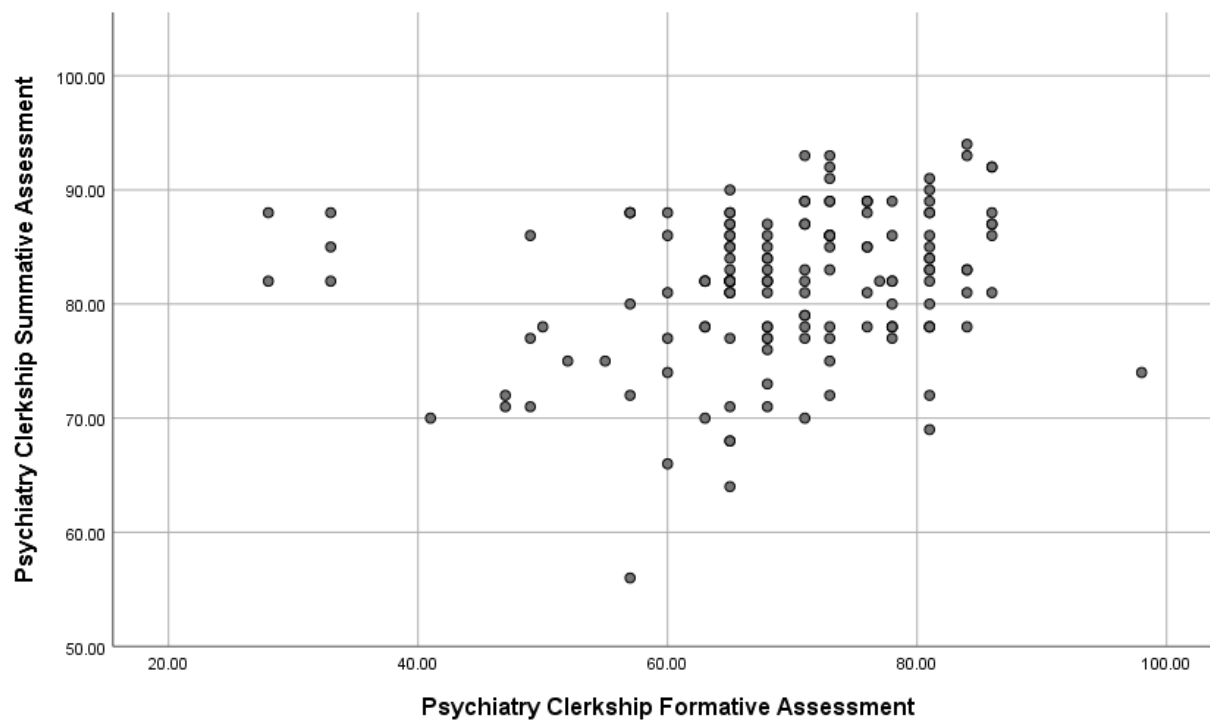
Data Set 3. Neurology Clerkship scatter plot. This data set illustrates the relationship between the Neurology Clerkship formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.



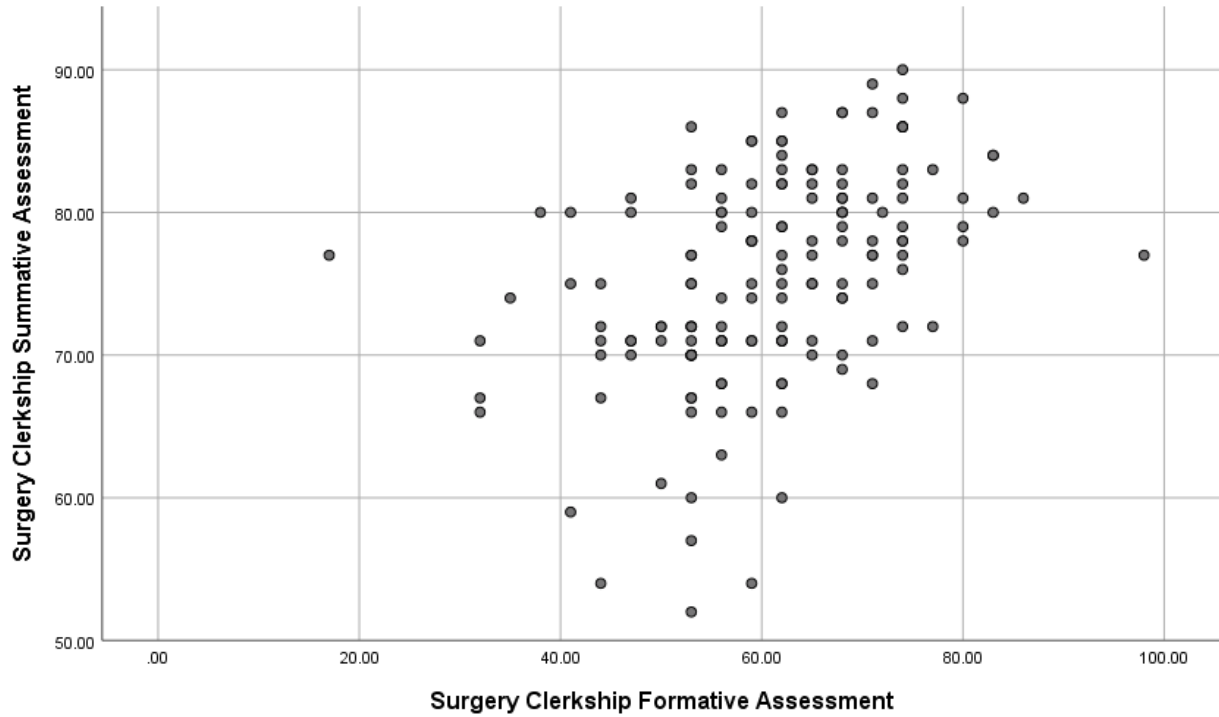
Data Set 4. OB-GYN Clerkship scatter plot. This data set illustrates the relationship between the OB-GYN Clerkship formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.



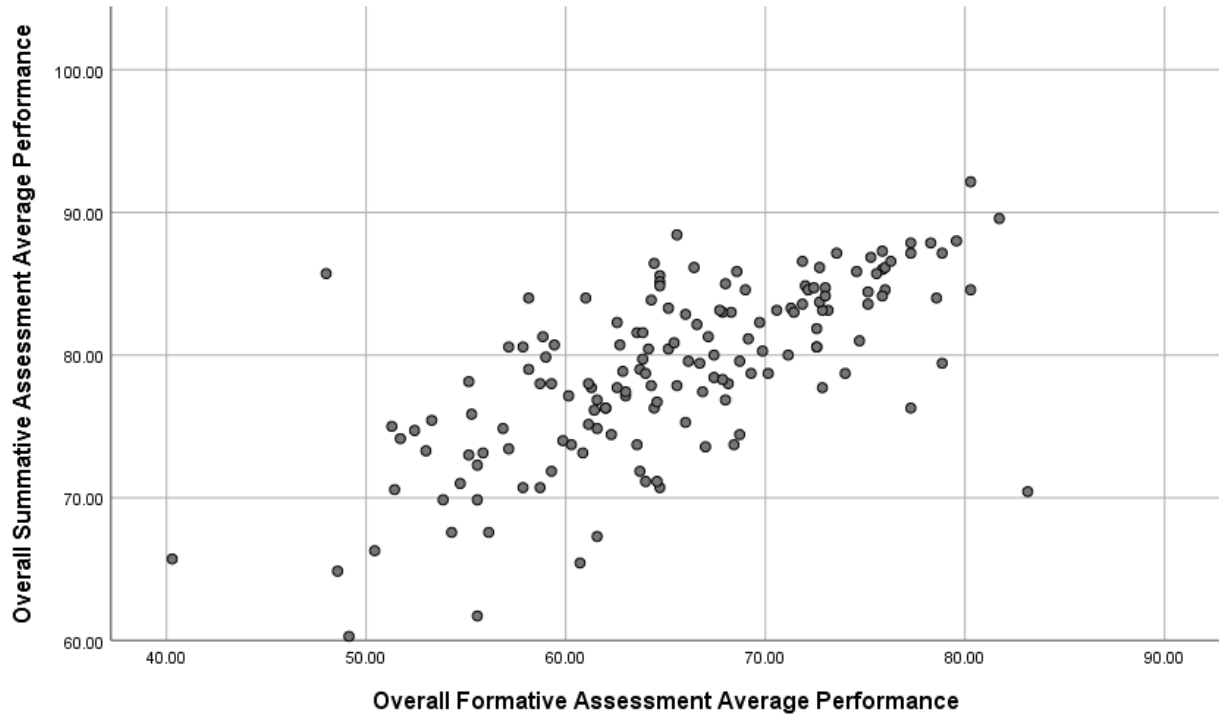
Data Set 5. Pediatrics Clerkship scatter plot. This data set illustrates the relationship between the Pediatrics Clerkship formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.



Data Set 6. Psychiatry Clerkship scatter plot. This data set illustrates the relationship between the Psychiatry Clerkship formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.



Data Set 7. Surgery Clerkship scatter plot. This data set illustrates the relationship between the Surgery Clerkship formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.



Data Set 8. Overall average performance scatter plot. This data set illustrates the relationship between the overall average formative and summative assessment performance by third-year undergraduate medical students at Northeastern Medical School's graduating classes of 2019 and 2020.