PENNSYLVANIA EDUCATIONAL LEADERSHIP

VOLUME 41 - NUMBER 1
SPRING / SUMMER - 2022

Pennsylvania ASCD
LEARN • TEACH • LEAD

PENNSYLVANIA EDUCATIONAL LEADERSHIP

VOLUME 41 | NUMBER 1 | SPRING/SUMMER 2022

EDITOR
Mary A. Wolf
PennWest University

REVIEWERS
Robin Fillman
Riverview Intermediate Unit

Karen Tyler
Edinboro University of Pennsylvania

Elaine Wilson
University of Pittsburgh at Johnstown

Felix Yerace
South Fayette School District

William Denny (Technical Reviewer)
California University of Pennsylvania

EDITOR EMERITUS
Robert F. Nicely, Jr.

LAYOUT DESIGNER
Sarah Smeltz
BLaST Intermediate Unit 17

Pennsylvania Educational Leadership is an official publication of the Pennsylvania Association for Supervision and Curriculum Development. Since 1996 the journal has received numerous awards from the international Association for Supervision and Curriculum Development. These awards include "best affiliate journal," "outstanding affiliate article" (multiple times), and "outstanding affiliate journal."

The views expressed or implied in the journal are not necessarily official positions of Pennsylvania ASCD. Membership in the Pennsylvania ASCD includes a subscription to *Pennsylvania Educational Leadership*, Pennsylvania ASCD *Update*, and selected occasional publications.

PENNSYLVANIA ASSOCIATION FOR SUPERVISION AND CURRICULUM DEVELOPMENT

PRESIDENT
Brian Griffith

PAST PRESIDENT
Barbara Davis

PRESIDENT-ELECT
Bryan O'Black

EXECUTIVE DIRECTOR
Deanna Mayers

WEBSITE
https://PASCD.org/

TABLE OF CONTENTS

VOLUME 41 | NUMBER 1 | SPRING/SUMMER 2022

RESEARCH MANUSCRIPTS

CALL FOR SUBMISSIONS

Pennsylvania Educational Leadership (PEL) is an open-access, peer-reviewed journal published twice per year and accepts manuscripts year-round. Topics address the interests and concerns of Pennsylvania educators at all levels. We welcome a wide variety of manuscripts including (but not limited to) single study inquiries, qualitative and quantitative research, theoretical and conceptual pieces, historical analyses, literature reviews, action research, and first-person narratives. Beginning spring 2014, the journal began including a Practitioner's Page highlighting the voices, thoughts, and opinions of educators in the field. Submissions for the Practitioners Page can take a variety of formats including (but not limited to) book reviews, policy reviews, and critical reflections on current educational issues and trends. Individuals choosing to submit to the Practitioners Page should follow the same submission guidelines as those submitting manuscripts with the exception of the Abstract. Authors must also indicate that the submission is intended for the Practitioners Page on the cover sheet.

Manuscripts should be emailed to Editor Mary Wolf (California University of Pennsylvania) at wolf@pennwest.edu for initial review. Submissions evaluated as appropriate for review are then sent to three readers for blind review. Manuscripts should follow the guidelines set forth by the American Psychological Association.

Before submitting a manuscript to PEL, please consider the following guidelines carefully:

- Your manuscript should be submitted as a single Word document and include a cover sheet, abstract, body/text, tables, charts, and figures (if applicable), and references list. If possible, please include the Digital Object Identifiers (DOI) for all electronic sources. The manuscript should be typed in 12-point font, Times New Roman, with one-inch margins. The text should be double-spaced.
- The cover sheet should include the title and author information, including contact information for the primary author, including mailing address, email address, and phone number. On this page, the author should indicate that the manuscript has not been submitted elsewhere for publication. If the manuscript involves the use of human subjects, the author should indicate whether Institutional Review Board approval has been granted unless deemed exempt.
- The second page of the submitted manuscript is the abstract page. The abstract should be 150 words or fewer. The abstract should include the purpose of the manuscript and essential findings or discussion points.
- The author(s) should remove any references that might be self-identifying from the body of the text to ensure blind review of the manuscript.
- The references page will follow the body of the text and any tables, charts, or figures. Please be sure to check that all in-text citations match references in the list and that the list is properly formatted using APA guidelines. Please include the DOI for electronic sources.
- The deadline for the fall/winter 2022 edition is August 15, 2022.
  Questions regarding a possible submission, submissions under review, or submissions requiring revision can be directed to Mary Wolf (Editor) at wolf@pennwest.edu.

LETTER FROM THE EDITOR

This past July, I moved into the role as editor of the Pennsylvania Leadership Journal. As a PASCD member for almost 30 years, I watched the journal evolve and marveled at the work of past editors. Most recently, Kathleen Provinzano led the process and provided mentorship as I moved from assistant editor to co-editor, and now editor. I thank Kathleen for her commitment and professionalism over the years.

Peer reviewers were recruited and have supported the process immensely. A special thanks to the reviewers. They stepped in on short notice and served well.

With so many educator professionals in the field conducting research to improve their knowledge, it is with pleasure that we highlight six studies. In the journal, the authors share their methods and outcomes with the readers of PEL.

In the first article, Joseph Jablonski examined the efficacy of a universally delivered social and emotional learning (SEL) program. He provides evidence through a quantitative analysis conducted through a quasi-experimental method, and results highlight the efforts and stress factors involved with implementing a new program during the COVID-19 pandemic.

In the second article, Brian Randall addresses the barriers to parental involvement for non-nuclear families; a growing sub-group in society. Through individual interviews, a focus group interview, and a writing prompt, school administrators shared their experiences regarding non-nuclear family barriers.

Grading and assessment reform is the next area of focus presented by Divonna Stebick. A variety of contemporary grading approaches have gained widespread popularity. A challenge is that school districts define these terms differently and models vary widely in their implementation. The researchers identified the common ground about what an effective system for grading and assessment should include.

Summer mathematics learning loss in elementary students was the focus of a study presented by Joseph Anthes. The study used experimental and quasi-experimental designs and found information related to math learning loss as well as the effect of reading programs on math outcomes for first and second grade students in summer learning programs.

Melinda Burchard identifies the support needed to prepare pre-service special educators to co-teach personal financial content. Researchers examined interactions between self-efficacy for interventions with outcomes of gains in knowledge, budgeting skills, and quality of lesson design for teaching personal finance content. Qualitative analysis revealed three common themes of importance for students to understand their personal financial journey.

In the final article, Amml Hussein examines the affective dimensions of the college sophomore year experience. Her study explored the relationship between social and emotional learning variables on student success and retention during the sophomore year of the undergraduate experience. An examination of the affective levels of second- and third-year students commenced as measured by responses to a situational judgment test. The study offers recommendations for student support services and retention initiatives.

I hope that you enjoy reading the research conducted by our Pennsylvania colleagues and find the impact of their outcomes useful and applicable. If you have research or best practices to share, please consider submitting your manuscript for consideration in the fall/winter publication.

Editor of the 2022 PEL Journal
wolf@pennwest.edu

PASCD PUBLICATION AWARD SPOTLIGHT

DR. EVA ALLEN

Dr. Eva Allen received the PASCD Outstanding Research and Publication Award in 2019. The Outstanding Research and Publication Award is presented to Pennsylvania ASCD member who has significantly impacted education through research and publication. Selection is based on the research design, conceptual framework, format, and publication. Her award was presented at the annual PASCD conference in Hershey, PA.

Dr. Allen's research investigated teacher perceptions of the influence of cultural care and invitational education (IE) on the formation of a positive teacher-student relationship with students of color in an urban elementary school. Cultural care is a theory of practice that utilizes a social-emotional approach for school improvement and to promote positive student outcomes. It is defined as a verbal or nonverbal gesture that displays a genuine interest in another person's social, emotional, mental, and physical well-being; simultaneously recognizing and acknowledging race and culture as a vital part of a person's identity. Cultural care must include respecting, valuing, and embracing culture from a value- and strengths-based perspective.

Conducted through qualitative participatory action research, this study examined teacher practices and perceptions in order to evaluate the influence of cultural care. The study utilized elements of the theoretical frameworks of IE, culturally relevant pedagogy, critical race theory, and self-efficacy theory. The findings were derived from analyses of pre- and post-intervention implementation, recorded observations and notes, and artifacts that were generated as a result of participation in a professional learning community that was focused on equity and care.

Findings indicated the importance of teachers listening to students with intentionality; recognizing students' basic and academic needs; and acknowledging students' presence, behavior, and growth, including making gestures of concern. Also, emergent in the findings was the significance of educators developing self-reflection and self-awareness as a part of practice, sharing personal experiences and stories, and engaging students in nonacademic conversations to facilitate positive relationships with them. One unexpected outcome concerned student-initiated conversations on race. The participant educators reported that students were comfortable in talking and asking questions about race-related topics that are often difficult to discuss.

In addition to her scholarship, her work and service have also shaped the field of education. She is currently the Learning Environment Specialist for the Pittsburgh Public School District. Dr. Allen serves as a member of the Ed.D. Educational Leadership Alumni Advisory Board for the Ed.D. Educational Leadership program at Duquesne. Currently, she collaborates with equity and diversity experts in the field to establish Culturally Connected Education, a professional development and consulting firm that assists individuals,

schools, and districts in meeting and achieving their goals.

Allen, E. J. (2017). Cultural care and inviting practices: Building relationships in an urban elementary school. Teacher perspectives in forming positive teacher-student relationships based on care and equity [Doctoral dissertation, Duquesne University]. https://dsc.duq.edu/etd/160

NOMINATE A PASCD MEMBER FOR THE 2022 OUTSTANDING RESEARCH AND PUBLICATION AWARD

The Outstanding Research and Publication Award is presented to Pennsylvania ASCD member who has researched and published within the past two years. The person being nominated shall submit the published article, book, book chapter, or research report with the application form. Submitted documentation will be reviewed on the basis of design, conceptual framework, format, and publication. The award application may be found at https://www.pascd.org/awards.

EXAMINING THE EFFICACY OF UNIVERSALLY DELIVERED SOCIAL AND EMOTIONAL LEARNING CURRICULUM, SECOND STEP: A QUANTITATIVE STUDY OF STUDENT BEHAVIOR DURING THE COVID-19 PANDEMIC

Joseph A. Jablonski, Ed.D
Eric J. Bieniek, Ph.D., BCBA-D
Rich Busi, Ph.D.
Edward Nientimp, Ed.D.

PENNSYLVANIA EDUCATIONAL LEADERSHIP | VOLUME 41, NUMBER 1

ABSTRACT

This study examined the efficacy of *Second Step*, a universally delivered social emotional learning (SEL) program. Evidence is supplied through a quantitative analysis conducted through a quasi-experimental method. The method utilizes a McNemar test to determine the statistical significance of negative behavior prevalence in two school years with a sample of 505 students. Through this study, it was determined that *Second Step* had a statistically significant effect on the behaviors of students. While teachers felt the stress of teaching a newly implemented SEL curriculum during the COVID-19 pandemic, results from the study highlight the overall benefits of implementing the program.

INTRODUCTION

Public schools have an obligation to ensure that their students have an opportunity to receive a free and appropriate public education regardless of exceptionality (Individuals with Disabilities Education Act, 2004). The obligations imposed on public schools in this regard emanate from federal law and are also firmly rooted in the traditions of the American educational system. However, what if a student's social or emotional needs preclude him or her from attaining that education? What if a student has social and emotional needs, but they do not impact their ability to learn? Do school districts have an obligation to meet the social and emotional needs of all their students? Current trends in school systems, along with legislative efforts, have worked in recent years to provide the answer to these questions.

Over the past 20 years, the concept of addressing the social and emotional needs of students has been brought to the forefront of educational policy and practice (Weissberg et al., 2015). In 1994, a group of educators, researchers, and advocates met to discuss the needs of enhancing students' competence of social and emotional constructs, academic performance, health, and citizenship. To address these needs, they developed the conceptual framework of "social and emotional learning." This framework of social and emotional learning (SEL) sought to prevent mental health and behavioral problems by teaching students the competence of five domains: self-awareness, self-management, responsible decision making, relationship skills, and social awareness (Weissberg et al., 2015). This culminated in the establishment of the Collaborative for Academic, Social, and Emotional Learning (CASEL). The mission of CASEL is to establish evidence based SEL that focuses on the five identified domains, from preschool through high school, in educational systems throughout the country (Weissberg et al., 2015).

The purpose of this study was to examine a school's implementation of an evidence based SEL program, *Second Step*, and quantify its efficacy through the extent and amount of externalizing behavioral infractions as compared to the previous school year.

LITERATURE REVIEW

Since the establishment of SEL as an instructional framework, there has been a growing body of evidence demonstrating the efficacy of SEL in establishing prosocial behaviors and emotional competence in students. Studies have shown the impact SEL has on students' prosocial and emotional competence in different settings and programs (e.g., Bierman et al., 2010; Chi-Ming et al., 2004; Doughty, 1997; Low et al., 2015).

As the concept of SEL is being addressed in the research and school-based sector, it is also paralleling the mental health legislation occurring in the public sector. For instance, in 1990, the state of Pennsylvania began to address the issue of young adults abusing drugs, alcohol, and/or tobacco. In a concerted effort, Act 211 was enacted, and section 1547 of the Pa School Code was added. This legislation mandated that all school districts must implement a drug, alcohol, and tobacco prevention program. Through this program, the Student Assistance Program (SAP) was developed (Act 211, 1990). SAP helped to screen and find students who may be at risk and provide intervention and/or counseling services to them. This system was found to be successful and, as a result, in 2006, Chapter 12 of the Pennsylvania School Code was expanded to increase the scope of SAP to include a mental health component (Student Services, 2006). More legislation regarding student mental health was to follow. Specifically, in 2019, Pennsylvania enacted Act 18, which required schools to recognize the impact of trauma on students and provide them with the necessary support (Act 18, 2019). SEL is an integral part of trauma-informed care, as it fosters resilience capabilities through emotional literacy and problem-solving (Payton et al., 2008).

Aside from SEL increasing prosocial and emotional behaviors and working through the lens of trauma-informed education, there is an academic component of SEL that justifies its inclusion as an educational practice in schools. Historically, Abraham Maslow made the argument that for higher-level thinking to occur, a human must first receive their basic and psychological needs. He created a hierarchy with psychological necessities following basic needs, before self-fulfillment (Maslow, 1943). From a very rudimentary standpoint, if we are not addressing basic needs, based on Maslow's theory, motivation for academic achievement can never be met. A student's emotional needs will supersede their academic needs due to their emotional state (Plumb et al., 2016). Therefore, if schools want to ensure academic growth and achievement, they must first meet the student's fundamental social and emotional needs. To further substantiate SEL's effect on academic performance, a recent meta-analysis of hundreds of these studies was conducted, demonstrating a positive correlation between social/emotional well-being and higher academic achievement in participants of social and emotional learning programs (Durlak et al., 2011; Taylor et al., 2017). There are also potential financial benefits to school districts when SEL is prioritized. In 2015, Belfield et al. conducted a benefit-cost analysis (BCA) of various SEL programming. Through their analysis, the researchers found that the benefit of SEL instruction substantially outweighed the cost, where in some cases a $1 expenditure translated to an $11 benefit (Belfield, et al., 2015).

Attending to students' social and emotional needs is immensely important. Moreover, research has also shown economically disadvantaged students are more likely to experience maladaptive behaviors. In a study of 327,617 participants, students who came from lower-income families were found to experience more negative externalizing and internalizing behaviors (Korous et al., 2018). Other research has found that low socioeconomic status negatively correlates to higher conduct problems in students of grades K, 1, 2, and 3

(Dodge & Pettit, 1994).

Therefore, the purpose of this study is to evaluate a program designed to improve SEL for elementary students within a setting of high numbers of economically disadvantaged students and determine whether there was any effect on the overall behavior of students. This will be determined through the comparison of a school's Office Disciplinary Referrals (ODRs) from the previous year.

METHODOLOGY

The school selected for this study was chosen because it has seen a steady increase in economically disadvantaged students in recent years. This school is one of five elementary schools located within a school district in northwestern Pennsylvania with a district enrollment of approximately 6,500 students. Of the 630 K-5 students attending the school being studied, 49% are considered to be economically disadvantaged and approximately 8% are identified as English Language Learners (ELL). Another 12% of the student population receive Special Education services and 26 of those students are identified as needing Emotional Support (ES) Services, accounting for 45% of the special education demographic (PA Future Ready Index, 2020).
The school's demographic has changed substantially over the past six years due in part to the consolidation efforts, as well as the addition of federally subsidized Section 8 housing within the school's boundaries. Prior to the consolidation, in the school year (SY) 2012-2013, the school had an enrollment of 561 students. With this enrollment, 25.6% of students were considered economically disadvantaged. After the consolidation, in SY 2013-2014, the school's enrollment increased 36% to 763 students. The increase in student enrollment correlated with an increased percentage of economically disadvantaged students at a rate of 35.5%. Moreover, there was the establishment of federally subsidized housing within the school's boundaries. Due to this, within four years, the economically disadvantaged percentage grew substantially and reached 48.5% in SY 2018-2019. In conclusion, in the span of six years, the school saw its economically disadvantaged percentage almost double in size (PA Future Ready Index, 2020).

With the influx of economically disadvantaged students, there was also an increase in the number of students with socially inappropriate and maladaptive emotional behaviors. This increase of inappropriate and maladaptive social and emotional behaviors led to reflective questioning for administrators. Prior to this study, the school and its district did not have an SEL curriculum that provided instruction in these areas to students, instead a positive behavior intervention and supports (PBIS) framework was utilized. However even with the PBIS framework, the school found itself without the necessary tools to combat the increasing negative social and emotional behaviors with no systematic, explicit, evidence-based instruction to educate the students on their deficiencies and *Second Step* SEL program was identified as a possible solution to help address these areas.

Second Step is an SEL program that provides age-appropriate differentiated K-5 lessons that focus on building the five core competencies of SEL with defined plans for each grade level that are CASEL, PBIS, MTSS/RTI, and Common Core Aligned. There are 22-25 lessons that are designed to be delivered weekly. Aspects of the lessons include direct instruction, modeling, reflective thinking, discussion, and group activities. These lessons are centered around identifying and naming feelings, empathy, emotion management, and problem solving. The program also provides embedded integration for existing PBIS systems and academic courses as well as communication materials for home (CASEL, 2013; Jones et al., 2017, Low et al., 2019; Low et

al., 2015).

The efficacy of the program is examined by answering the following research question: What is the effectiveness of the *Second Step* SEL program on student behavior as evidenced by Office Disciplinary Referrals (ODRs) when it is delivered in the regular education classroom at a universal level?

DESCRIPTION OF PROCEDURES

The *Second Step* program was implemented in the 2020-2021 school year at the study school. Program implementation started at the opening teacher in-service, where initial professional development was provided to teachers. Throughout the school year, homeroom teachers provided weekly lessons to both regular and special education students in their classrooms. *Second Step* has 22 weekly lesson plans for grades 1-5 and 25 weekly lessons for kindergarten. The start and end guidelines, whole school assembly, and whole school announcement schedule are included in Appendix A.

Since implementation occurred with flexibility, meaning that it was delivered at different times by different teachers, this variance was discussed with the program developers. An interview was conducted with the Vice President of Education and Research at *Second Step* on January 23, 2020. In the interview, it was confirmed that implementation fidelity would not be adversely affected by integrating *Second Step* flexibly. The program was developed in this manner to allow the program to fit various school needs (T. Kim, personal communication, January 23, 2020).

At the conclusion of implementation, office discipline referral (ODR) data was collected and compared between the 2020-2021school year (with *Second Step*) to the 2019-2020 school year (no *Second Step*), and statistical analyses were completed to determine if differences existed.

RESULTS

Quantitative Office Disciplinary Referral Results-Effect on Overall Student Behavior

To analyze the effect of *Second Step* on student behavior, a pre/post analysis of Office Disciplinary Referrals (ODRs) was conducted. This analysis was conducted to answer the first research question: *What is the effectiveness of the Second Step SEL program on student behavior as evidenced by Office Disciplinary Referrals when it is delivered in the regular classroom at a universal level?*

To retrieve this information, two ad hoc reports were created. The first ad hoc report generated a list of students enrolled in grades Kindergarten through 4th grade during the 2019-2020 school year at the study school. The report gave each student a unique 32-digit alphanumeric character and nominally attributed a "Y" or "N" to each student depending on whether the student received an ODR during the first 128 days of the 2019-2020 school year. Multiple incidents conducted by a single student were not included to prevent the skewing of data attributed to anomalous individuals. The time-period of 128 days was used as the defined range because that was the length of time the study school was in session prior to the COVID-19 shutdown in 2019-2020.

The second ad hoc report was generated after the presentation of the independent variable, *Second Step*, in the 2020-2021 school year. The second report provided the same information as the first, over the same range of 128 days. The report included students in grades 1-5, as these students were in the cohort from

the prior year and matriculated to the next grade. After both the 2019-2020 and the 2020-2021 reports were created, they were cross-referenced and any student that was not on both lists was excluded from the sample set because they were not present in both pre/post-treatment.

Student Sample

The total number of students in the sample data who met the above criterion from the study school was 505. The students in the sample data came from the following cohorts:

- 107 students from kindergarten (19-20) to 1st Grade (20-21)
- 102 students from 1st Grade (19-20) to 2nd Grade (20-21)
- 105 students from 2nd Grade (19-20) to 3rd Grade (20-21)
- 95 students from 3rd Grade (19-20) to 4th Grade (20-21)
- 96 students from 4th Grade (19-20) to 5th Grade (20-21)

ODR Data and Analysis

Table 1 illustrates the total number of individuals who received ODRs in each cohort during the two school years. Overall, there was a 35.37% decrease in the total amount of ODRs for the student sample data. Kindergarten/1st had the largest decrease with 88% less than the previous year, followed by 4th/5th with 26.67%, 3rd/4th with 18.19%, and 1st/2nd with 11.12%. The cohort with the largest total amount of ODRs in 19-20 was 4th/5th with 30. This was followed by kindergarten/1st with 25. In 20-21, the largest total number of ODRs was 4th/5th with a total of 22, with the next highest being 2nd/3rd with 11.

Table 1

ODR Totals by Cohort

COHORT	STUDENTS	19-20	20-21	%+/-*
K/1st	107	25	3	-88%
1st/2nd	102	9	8	-11.22%
2nd/3rd	105	7	11	+36.36%
3rd/4th	95	11	9	-18.19%
4th/5th	96	30	22	-26.67%
TOTAL	505	82	53	-35.37%

Note. *Percent increase or decrease, year over year.

Four out of five cohorts experienced a decrease in the amount of ODRs. The exception was the 2nd/3rd cohort. That cohort experienced an increase of ODRs by 36.36%.

Statistical Significance

To determine the statistical significance of these increases/decreases in ODRs between the school years, a repeated measures McNemar statistical analysis test was completed to determine statistical significance in the difference in number of ODRs between school years. The McNemar test was chosen for its ability to analyze repeated measures specifically on non-parametric data.

The McNemar test was completed with an $\alpha = .05$ to determine statistical significance. The null hypothesis posited that after treatment with the Second Step program, there will be no change in behavior as evidenced by ODRs. The occurrence of behavioral incidents is delineated by a "Y" in Table 2, where an "N" indicated that there was no behavioral incident. Results of the crosstabulation statistical comparison are included in Table 2.

Table 2

McNemar Crosstabulation of Overall ODRs

			AFTER SECOND STEP		
			N	Y	TOTAL
Before *Second Step*	N	COUNT	369	27	423
	Y	COUNT	56	26	82
TOTAL		COUNT	452	53	505

As Table 3 illustrates, the p-value of .002 represents a significant statistical significance and subsequently a rejection of the null hypothesis. The alternative hypothesis that behavioral incidents were not equal to those prior to the *Second Step* program was accepted.

Table 3

McNemar Statistical Significance of Overall ODRs

	VALUE	EXACT SIG. (2-SIDED)
McNemar Test		.002[a]
N of Valid Cases	505	

Note. [a.] Binomial distribution used.

An additional McNemar statistical analysis was completed to determine if there was a level of statistical significance that was different at other grade levels. This analysis was completed with an **α**= .05 to determine statistical significance and the results of the crosstabulation are included in Table 4.

Table 4

McNemar Crosstabulation of Individual Grade Level ODRs

			AFTER SECOND STEP		
Grade level	Before *Second Step*		N	Y	TOTAL
Kindergarten/1st		N	81	1	82
		Y	23	2	25
	TOTAL		104	3	107
1st/2nd		N	88	5	93
		Y	6	3	9
	TOTAL		94	8	102
2nd/3rd		N	87	11	98
		Y	7	0	7
	TOTAL		94	11	105
3rd/4th		N	78	6	84
		Y	8	3	11
	TOTAL		86	9	95
4th/5th		N	62	4	66
		Y	12	18	30
	TOTAL		74	22	96
TOTAL		N	396	27	423
		Y	56	26	82
	TOTAL		452	53	505

The p-values used to quantify the amount of statistical significance are included in Table 5. There is a discrepancy in the p-values for the individual grade levels. Whereas the original null hypothesis can be rejected in the kindergarten/1st cohort, the null hypothesis in the 1st/2nd, 2nd/3rd, 3rd/4th, and 4th/5th cohorts remains, as the p-values were not statistically significant for these grade-level cohorts. Of note is that the p-value is so significant in the kindergarten/1st that it offsets the p-values of the four other cohorts in the statistical comparison of the whole school.

Table 5

McNemar Statistical Significance of Individual Grade Levels

Grade Level		Value	Exact Sig. (2-sided)
Kindergarten/1st	McNemar Test		<.001 [a]
	N of Valid Cases	107	
1st/2nd	McNemar Test		1.000[a]
	N of Valid Cases	102	
2nd/3rd	McNemar Test		.481[a]
	N of Valid Cases	105	
3rd/4th	McNemar Test		.791[a]
	N of Valid Cases	95	
4th/5th	McNemar Test		.077[a]
	N of Valid Cases	96	
	McNemar Test		.002[a]
TOTAL	McNemar Test	505	

Note. [a.] Binomial distribution used.

The data from this objective quantitative analysis suggests that the *Second Step* program had a statistically significant positive impact on lowering the occurrences of negative student behavior as evidenced by a decrease in the number of students having ODRs.

DISCUSSION OF STUDY AND RESULTS

Effect on School-Wide Behavior

Several studies have been conducted to provide an evidence base that supports the positive behavioral effect of students participating in various SEL programs (Durlak et al., 2011; Domitrovich et al., 2017; Eddy et al., 2003; Taylor et al., 2017). To quantitatively examine the behavioral effect of the *Second Step* program, the following research question was asked: *What is the effectiveness of the Second Step SEL program on student behavior as evidenced by Office Disciplinary Referrals when it is delivered in the regular education classroom at a universal level?*

This question was answered through the McNemar analysis which illustrated a statistically significant effect in the overall reduction of ODRs. This significant reduction provided evidence to support the *Second Step* program in realizing positive effects and impact on overall student behavior.

Effect on Grade Level Cohort Behavior

While there was a statistically significant reduction in school-wide ODRs, further cohort analysis provided a discrepancy in the reduction of ODRs within separate grade-level cohorts. The analysis found a significant statistical reduction in ODRs in the Kindergarten/1st and 4th/5th cohorts; however, the 1st/2nd, 2nd/3rd, and 3rd/4th cohorts did not have a statistically significant reduction. The explanation of this can be realized when examining the raw ODR data.

Upon further examination, a postulation can be made that since the kindergarten/1st and 4th/5th cohorts had substantially higher amounts of ODRs prior to the Second Step program, they were most likely to realize the greatest impact after participation in the program. Conversely, the other three cohorts had substantially lower amounts of ODRs pretreatment with Second Step, thereby leading to a negligible effect. These results lead to the hypothesis that because the kindergarten/1st and 4th/5th cohorts had larger ODR numbers before the program, those cohorts realized the greatest effect.

This hypothesis is on par with the results of a previous study by Jones et al., which concluded that SEL programs have a greater impact on students with lower social-emotional, and behavioral pretest scores (Jones et al., 2012). Also, research found that SEL programs can foster improved behavioral trajectories when administered at the universal level (Duncan et al., 2017). Further analysis of these cohorts would be recommended to see if the relative pattern of effectiveness continues longitudinally or if the results are attributed to a specific student population in those cohorts.

COVID-19 IMPLEMENTATION CONSIDERATIONS

As with all facets of life, COVID-19 has undoubtedly affected the education in general and specifically, the implementation of this program evaluation. It is worth noting that the study school utilized several instructional models throughout the school year. Roughly the first half of the year was conducted in a hybrid format where students alternated days the week for attendance with Monday being a virtual day. The second half of the year utilized in-person instruction with a virtual Monday. The combination of in-person, virtual, and hybrid instruction were utilized and thus, the *Second Step* program was disseminated through these various modalities. Despite the potential barriers that this could have presented, the positive feedback from

multiple data points supports *Second Step* as an effective SEL program regardless of the instructional model.

In speaking with program authors for the *Second Step* program, prior to implementation to ensure treatment fidelity throughout the implementation, they confirmed that the program was designed to be flexible when being delivered. This way the program could be tailored to meet the specific needs of the school (T. Kim, personal communication, January 23, 2020). The evidence from this study provides evidence to support the company's statements, despite the unique challenges that implementors faced over the 2020-21 school year. The ramifications of these findings could have a significant impact on schools as they look to utilize effective curriculums that can be delivered with flexibility. The ability to be able to provide instruction across modalities is important as schools find themselves adapting to changes in light of the COVID-19 pandemic. The practical results of this study and the effects of *Second Step* while being delivered both virtually and in-person, highlight the strength of the program.

CONSIDERATIONS/LIMITATIONS

This study was conducted during an atypical school year. In the 2020-2021 school year, the study school was operating under health and safety mitigation protocols due to the global COVID-19 pandemic. As a result, the school utilized various instructional modalities, however; the same behavioral expectations and reporting methods that were utilized in previous school years continued, even when students were participating across the three modalities. While this could have affected some of the overall ODR numbers, the additional information through instructor feedback provided evidence to support the reduction of ODR numbers because of the *Second Step* program.

An additional consideration is that the primary researcher was also an administrator at the study school. The school's PBIS team, school counselor, and emotional support teachers also took key roles in implementation as they filmed virtual assemblies, created posters, disseminated materials, developed skits, etc. It was reiterated multiple times that program implementation and evaluation was both anonymous and voluntary and there would be no repercussions for lack of participation. Implementors were provided with information to contact the partnering university's Institutional Review Board if they felt in any way compromised throughout this process.

Lastly, teacher effectiveness as it relates to the fidelity of instruction and integrity of the program was not evaluated. Therefore, the consistent level of instruction throughout the various classrooms could have varied.

SIGNIFICANCE AND RESEARCH CONTRIBUTIONS

The execution of the current study has resulted in a contribution to the evidence base surrounding SEL and the efficacy of the *Second Step* program. Another noteworthy outcome of this study is the evidence that was realized due to the study being conducted during a global pandemic. The effect and impact that is evidenced by this study bolster the recommendations of the reports that came out in advocacy for SEL as a critical component to reopening schools in the 2020-2021 school year following the COVID-19 school closures during the 2019-2020 school year (CASEL, 2020; Pennsylvania Back to School Task Force, 2020).

RECOMMENDATIONS FOR FURTHER RESEARCH

The findings of this study, coupled with previous research, provide a catalyst for further exploration and examination. A larger amount of information would be provided if the study was examined longitudinally. Through a multi-year analysis, evidence could be collected to determine if the different instructional modalities that were utilized due to COVID-19 had an impact on the overall ODR numbers. The longitudinal data could also be used to examine individual cohorts to determine if the relative level of difference in ODRs/effectiveness trended similarly. This data could then be analyzed through the same McNemar test to determine overall statistical significance, as well as significance on cohort subgroups. Research protocols could also be revised to determine if there are any potential cumulative effects of the rate of ODRs on subgroups.

A further longitudinal study could focus more qualitatively on identifying further contextual information from teachers regarding their attitudes and perceptions about SEL, its implementation, and the *Second Step* program. Teachers would be able to provide qualitative feedback on several inquiries including the quality of implementation in the second year, the level of comfort in instructing during the second year, the level of the behavioral effect, and if the added stress due to teaching a new curriculum during the pandemic was alleviated.

An additional extension of the study that could provide insight regarding the program would be to conduct a comparative analysis with the study school and the other elementary schools locally, regionally, or nationally.

CONCLUSION

The purpose of this study was to expand on the existing research surrounding SEL instruction in schools, specifically examining the Second Step program as one iteration of this curricular approach. By answering the research question posed, this study determined that the *Second Step* SEL program not only had a statistically significant impact on the reduction of behaviors as evidenced by the McNemar analysis. Results from the study suggest that overall, the positive benefits of implementing the program outweighed the negative, even when considering the delivery of this new curriculum during the COVID-19 pandemic.

The evidence that this study provided will help schools work to not only strengthen their students' social, emotional, and behavioral capacities but also help as they navigate their return to normalcy following the pandemic. As schools often function as a microcosm of their communities, programs like *Second Step* can be used to help encourage not only individual growth but growth as a school community. By using *Second Step* to help invest in the five core competencies of SEL; self-awareness, self-management, social awareness, relationship skills, and responsible decision making, schools will help students grow in their capacities and, in turn, have a greater impact on their communities.

REFERENCES

Act 18 of 2019: School safety package (changes to Act 44 and more). (2019, July 15). https://www.psba.org/2019/07/act-18-of-2019-school-safety-package-changes-to-act-44-and-more/

Act 211, 42 USCS § 15801 (1990).

Adedokun, O. A., & Burgess, W. D. (2012). Analysis of paired dichotomous data: A gentle introduction to the McNemar test in SPSS. *Journal of MultiDisciplinary Evaluation*, 8(17), 125-131.

Belfield, C., Bowden, A., Klapp, A., Levin, H., Shand, R., & Zander, S. (2015). The economic value of social and emotional learning. *Journal of Benefit-Cost Analysis*, 6(3), 508-544.

Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a multiyear universal social–emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology*, 78(2), 156–168. https://doi.org/10.1037/a0018607

CASEL. (2013). *Effective social and emotional learning programs: preschool and elementary edition* https://drive.google.com/file/d/17WVff3ynFN9VwjyFR0Wxye8YyivJG_7v/view

CASEL. (2020). *Reunite, renew, and thrive: social and emotional learning (SEL) roadmap for reopening school.* https://casel.org/wp-content/uploads/2020/07/SEL-ROADMAP.pdf

Student Services, PA Chapter 12 § 12.41 (2006)

Kam, C. M., Greenberg, M. T., & Kusché, C. A. (2004). Sustained effects of the PATHS curriculum on the social and psychological adjustment of children in special education. *Journal of Emotional & Behavioral Disorders*, 12(2), 66–78 https://doi.org/10.1177/10634266040120020101

Dodge, K. A., & Pettit, G. S. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development*, 65(2), 649–665. https://doi.org/10.2307/1131407

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development*, 88(2), 408–416. https://doi.org/10.1111/cdev.12739

Doughty, J. E. (1997). *The effect of a social skills' curriculum on student performance.* (ED 412260). ERIC. https://eric.ed.gov/?id=ED412260

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Duncan, R., Washburn, I. J., Lewis, K. M., Bavarian, N., DuBois, D. L., Acock, A. C. & Flay, B. R. (2017). Can universal SEL programs benefit universally? Effects of the positive action program on multiple trajectories of social-emotional and misconduct behaviors. *Prevention Science: The Official Journal of The Society for Prevention Research*, 18(2), 214–224. https://doi.org/10.1007/s11121-016-0745-1

Eddy, J. M., Reid, J. B., Stoolmiller, M., & Fetrow, R. A. (2003). Outcomes during middle school for an elementary school-based preventive intervention for conduct problems: Follow-up results from a randomized trial. *Behavior Therapy*, 34(4), 535–552.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Jones, S. M., Bouffard, S. M., & Society for Research in Child Development. (2012). Social and emotional learning in schools: From programs to strategies. Social policy report. Volume 26, Number 4. *In Society for Research in Child Development.*

Jones, S. M., Brush, K., Bailey, R., Brion-Meisels, G., McIntyre, J., Kahn, J., & Nelson, B. (2017). *Navigating SEL from the inside out: Looking inside & across 25 leading SEL programs: A practical resource for schools and OST providers: (Elementary school focus).* Retrieved from http://www.wallacefoundation.org/knowledge-center/Documents/Navigating-Social-and-Emotional-Learning-from-the-Inside-Out.pdf.

Korous, K. M., Causadias, J. M., Bradley, R. H., & Luthar, S. S. (2018). Unpacking the link between socioeconomic status and behavior problems: A second-order meta-analysis. *Development and Psychopathology*, 30(5), 1889–1906. https://doi.org/10.1017/S0954579418001141

Low, S., Cook, C. R., Smolkowski, K., & Buntain-Ricklefs, J. (2015). Promoting social-emotional competence: An evaluation of the elementary version of Second Step. *Journal of School Psychology*, 53(6), 463–477. https://doi.org/10.1016/j.jsp.2015.09.002

Low, S., Smolkowski, K., Cook, C., & Desfosses, D. (2019). Two-year impact of a universal social-emotional learning curriculum: Group differences from developmentally sensitive trends over time. *Developmental Psychology*, 55(2), 415–433. https://doi.org/10.1037/dev0000621.supp

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50, 370-396. http://psychclassics.yorku.ca/Maslow/motivation.htm

PA Future Ready Index. (2020). *School fast facts*. https://futurereadypa.org/School/FastFacts?id=202197139056179021002009048154119163179212173136

Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., Pachan, M., & Collaborative for Academic, Social and Emotional Learning (2008). The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews. (Technical Report). *In Collaborative for Academic, Social, and Emotional Learning.* Retrieved

from https://www.casel.org/wp-content/uploads/2016/08/PDF-4-the-positive-impact-of-social-and-emotional-learning-for-kindergarten-to-eighth-grade-students-executive-summary.pdf

Pennsylvania Back to School Task Force. (2020). *Pennsylvania school reopening task force report.* https://www.pasa-net.org/Files/Coronavirus/SchoolReopeningReport.pdf

IEP, Pennsylvania School Code 22 Pa Code § 14.131. (2008).

Plumb, J. L., Bush, K. A., & Kersevich, S. E. (2016). Trauma-sensitive schools: An evidence-based approach. *School Social Work Journal*, 40(2), 37–60.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, 88(4), 1156–1171. https://doi.org/10.1111/cdev.12864

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., & Gullotta, T. P. (2015). Social and emotional learning: Past, present, and future. In R. P. Weissberg, J. A. Durlak, C. E. Domitrovich, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice.* (pp. 3–19). The Guilford Press.

APPENDIX
SECOND STEP IMPLEMENTATION TIMELINE
SY 2020-2021

- Summer Implementation completed on an on-demand basis prior to the start of school on 8-31-20
    - o District introductory video to staff
    - o Administrator 2 hr. online Second Step training
    - o Administrator onboarding video conference
    - o Teachers 2 hr. online Second Step training
    - o Parent introduction posted with Welcome Back Letter
    - o Parent introduction included in Welcome Back Video
- August
    - o Teacher Inservice
        - - Second Step Orientation Meeting 1
- September
    - o Began 5-week Second Step COVID-19 Community Rebuilding Unit
    - o Weekly Staff-led PD at Team Meetings.
    - o Faculty Meeting
        - - Orientation Meeting 2
- October
    - o Month 1 K-5 Weekly Lessons w/ Home Link correspondence and daily reinforcement strategies. Weekly Staff-led PD at Team Meetings.
    - o Month 1 Kick-Off Staff Meeting
    - o Month 1 Virtual Kick-Off Assembly
    - o Weekly announcements highlighting the week's SEL focus
    - o Monthly parent SEL update posted to school website
- November
    - o Month 2 K-5 Weekly Lessons w/ Home Link correspondence and daily reinforcement strategies. Weekly Staff-led PD at Team Meetings.
    - o Month 2 Kick-Off Staff Meeting
    - o Month 2 Virtual Kick-Off Assembly
    - o Weekly announcements highlighting the week's SEL focus
    - o Monthly parent SEL update posted to school website
- December
    - o Month 3 K-5 Weekly Lessons w/ Home Link correspondence and daily reinforcement strategies. Weekly Staff-led PD at Team Meetings.
    - o Month 3 Kick-Off Staff Meeting

- o Month 3 Virtual Kick-Off Assembly
- o Weekly announcements highlighting the week's SEL focus
- o Monthly parent SEL update posted to school website

• January

- o Month 4 K-5 Weekly Lessons w/ Home Link correspondence and daily reinforcement strategies. Weekly Staff-led PD at Team Meetings.
- o Month 4 Kick-Off Staff Meeting
- o Month 4 Virtual Kick-Off Assembly
- o Weekly announcements highlighting the week's SEL focus
- o Monthly parent SEL update posted to school website

• February

- o Month 5 Kick-Off Staff Meeting
- o Month 5 Virtual Kick-Off Assembly
- o Month 5 K-5 Weekly Lessons w/ Home Link correspondence and daily reinforcement strategies. Weekly Staff-led PD at Team Meetings.
- o Weekly announcements highlighting the week's SEL focus
- o Monthly parent SEL update posted to school website

• March

- o Month 6 Kick-Off Staff Meeting
- o Month 6 K-5 Weekly Lessons w/ Home Link correspondence and daily reinforcement strategies. Weekly Staff-led PD at Team Meetings.
- o Month 6 Virtual Kick-Off Assembly
- o Weekly announcements highlighting the week's SEL focus
- o Monthly parent SEL update posted to school website

• April

- o Completion of Second Step Program
- o Continue to reinforce strategies that were taught/learned

ABOUT THE AUTHOR

Joseph Jablonski, Ed.D., is the Principal of Grandview Elementary in the Millcreek Township School District. His interests include social emotional learning, positive behavior intervention supports, and trauma-informed practices.
jablonski@mtsd.org

NON-NUCLEAR FAMILIES:
ADDRESSING A GROWING SUBGROUP FROM THE SCHOOL ADMINISTRATOR'S PERSPECTIVE

Brian A. Randall, Ed.D.

PENNSYLVANIA EDUCATIONAL LEADERSHIP | VOLUME 41, NUMBER 1

ABSTRACT

The purpose of this transcendental phenomenological study was to understand the shared lived experiences of Central Pennsylvanian high school administrators' efforts to address the barriers to parental investment for non-nuclear families. The theory guiding this study was the path-goal theory, which builds on motivational theory to select the proper leadership style to achieve the desired outcome. The sites of this study included 12 school districts, which varied in size. Participants included different Central Pennsylvania school district administrators, purposefully sampled by selecting those who were knowledgeable about the phenomenon to share their experiences. Through individual interviews, a focus group interview, and a writing prompt, school administrators were encouraged to share their experiences regarding non-nuclear family barriers. Through participants experiences, it was established that school administrators put forth extra effort to assist non-nuclear families, a growing sub-group in American society.

Keywords: family structure, leadership practices, non-nuclear, parental investment

INTRODUCTION

"I think awareness is important. . . remind administrators that, you know, this may be a different subgroup too. . . understand that they may have different needs than your other parents. And how can we support those needs?" This comment, made by a rural principal, introduces the idea that non-nuclear families, a diverse family setting that does not include two biological parents within the same household (Bengtson, 2001; Miller, 2020), may need to be considered their own subgroup to help educators ensure they are providing an equitable education to non-nuclear families. Non-nuclear families face barriers such as time and resource constraints, cultural differences, and a lack of familiarity with the school system (Hamlin & Flessa, 2018) that could inhibit their ability to invest in their child's education. Family structure and its impact on educational outcomes are well-researched from the perspective of the family (Cavanagh & Fomby, 2019; Sun & Li, 2011); however, little research is available on school administrators' viewpoints and how they experience parental investment challenges of non-nuclear families. Family structure may impact academic outcomes (Santin & Sicilia, 2016), but it may also affect parental investment with the school (Garbacz et al., 2017; Myers & Myers, 2015), student attendance (Lim et al., 2019; Moehling, 2004), and student discipline (Gordon & Fefer, 2019). The purpose of this transcendental phenomenological study was to elucidate the experiences of high school administrators regarding the barriers to parental investment for non-nuclear families.

CONCEPTUAL FRAMEWORK

Path goal theory and expectancy theory provided the conceptual framework for this study. The synthesis of these two theories asserts that leadership styles are selected based on the subordinate's motivation. Subordinates in this study are parents raising students in a non-nuclear household. The first of four assumptions of their motivation for parental investment includes expectations about their needs, motivations, or past experiences (Razik & Swanson, 2010). The second assumption is that their behavior is a conscious choice that the parents are free to choose as they deem most appropriate (Lunenburg, 2011). Third, parents will want different things from the school (Vroom, 1964). The last assumption is that parents will choose the alternative that is most beneficial to themselves (Northouse, 2019). Based on their motivation, school administrators select a leadership style, either supportive, directive, achievement-oriented, or participative leadership, to best serve that parent and, therefore, the student in that situation (Olowoselu et al., 2019). This interaction between the parent's type of motivation for parental investment and the school administrator's corresponding leadership style helped describe the shared experiences of Central Pennsylvanian high school administrators' efforts to assist non-nuclear parents through parental investment barriers.

METHODS

Qualitative research serves to identify a holistic account of the phenomenon and interpretation of the meaning-making process (Patton, 2015). Through the qualitative research design, the researcher presented a holistic account of the school administrators' experiences to describe the complex interactions of multiple factors to understand the intricate picture of school administrator experiences regarding parental investment barriers of non-nuclear families. Phenomenological research describes a common meaning from lived experience for a group of individuals (Creswell & Poth, 2018). There is ample research on the non-nuclear

family structure and its impact on academic performance, but the child's education is more than academic success (An & Sorensen, 2017; Brock & Edmunds, 2010; Cavanagh & Fomby, 2019; Laursen et al., 2019). The research is limited in explaining the phenomenon of parental investment barriers from a school administrator's perspective. This qualitative study gave a voice to the school through school administrators to provide an in-depth look at their experiences with parental investment barriers of non-nuclear families.

PARTICIPANTS AND SETTINGS

This study was conducted across 12 sites in Central Pennsylvania. It included 4 small schools (less than 600 students), 4 medium schools (601-1,000 students), and 4 large schools (over 1,000 students). Collectively the sites had 64 administrators, but ultimately 12 participants were chosen for this study. Sites and participants were provided a pseudonym to ensure participant privacy (Table 1).

Table 1

Participants Demographics

PARTICIPANT	GENDER	SCHOOL	ROLE	YEARS OF EXPERIENCE HS/ADMIN	FOCUS GROUP PARTICIPANT
Donna Dunkled	Female	Monroe HS	Guidance Counselor	5/15	Yes
Tom Jennings	Male	Fillmore HS	Principal	18/3	Yes
Jasmine Russsel	Female	Monroe HS	Principal	11/2	No
Walter Boyd	Male	Jackson HS	Principal	24/12	No
Don Hoffman	Male	Jackson HS	Principal	27/13	Yes
Judy Caldwell	Female	Polk HS	Principal	33/9	Yes
John Black	Male	Taylor HS	Principal	14/6	No
Floyd Carter	Male	Martin HS	Principal	15/5	No
Dave Davis	Male	Madison HS	Principal	15/7	No
Marty Daniel	Male	Adams SD	Director of Education	16/6	No
Shawn Ross	Male	Tyler HS	Principal	16/7	No
Jake Yates	Male	Tyler HS	Assistant Principal	14/2	No

DATA COLLECTION

Data collection consisted of interrelated activities that served to gather good quality information to answer the following research questions (Creswell & Poth, 2018). The central research question for this study was: What is the essence of Central Pennsylvania school administrators' shared lived experiences regarding parental investment barriers for non-nuclear families? The study involved three data collection methods to triangulate the information and ensure trustworthiness while gathering, analyzing, and interpreting the data (Creswell & Poth, 2018). The methods that were implemented in this study were individual interviews, a focus

group interview, and a writing prompt. First, the researcher conducted an individual interview with each participant. Interviews lasted between 60 and 90 minutes and were conducted at locations that were convenient for the school administrator. A focus group was then scheduled with four participants who had unique experiences that informed the study. All interviews and the focus group were audio-recorded and transcribed verbatim. Participants were also asked to complete a writing prompt with three questions that garnered the participants' thoughts, beliefs, and views on the non-nuclear family structure and its implications on educational outcomes, their experiences with parental investment barriers, and their thoughts on leadership.

DATA ANALYSIS

Individual interview and focus group data was transcribed and then organized and coded using NVivo software. The researcher analyzed themes by looking for common phrasing and approaches to non-nuclear families. The researcher developed textural and structural descriptions of the individual interviews, focus group interview, and written prompts from the themes that developed through analysis.

After memoing significant statements, the researcher provided all data equal value and clustered data into emerging themes. Next, the researcher grouped significant statements into broader units of information called themes (Creswell & Poth, 2018). The researcher then constructed a list of significant statements that were relevant to the research question. The researcher utilized themes to develop a narrative on what was experienced by the participants and how they experienced the phenomena. This process allowed the researcher to understand the phenomenon through the participant's experiences using the transcripts verbatim. Finally, the essence of the phenomenon was created by using the textural and structural descriptions of Central Pennsylvania school administrators' perceptions regarding parental investment barriers for non-nuclear families.

FINDINGS

The general shared experiences of school administrators regarding parental investment barriers for non-nuclear families was that the experience is frustrating. All participants expressed a desire to do whatever it takes to help a child become successful. In non-nuclear households, barriers exist that inhibit parental investment. Through their leadership roles, the school administrators attempted to remove or diminish the obstacle the best they could. This would require additional work, time, and resources; all of which are limited at public schools. This was a source of frustration, not toward the non-nuclear household, but toward the educational system as a whole.

A major theme that emerged through this study was the administrators were faced with extra work when addressing parental investment barriers for non-nuclear families. The path-goal theory (House & Mitchell, 1975) entails the leader (administrator) clarifying a path around the obstacle which would require additional effort by the leader (administrator); therefore, finding this major theme supported the decision to use path-goal theory as a structure for this study. The subthemes that support this major theme of additional work were drawn from the participants' comments.

PARENTAL MOTIVATIONS AND THE SCHOOL'S ROLE

Participants expressed their level of understanding of parental expectations during the study. However, those expectations were dependent on the parent's investment. The parental investment barriers that non-nuclear families faced resulted in extra work for the school administrator. Shawn, the principal at Tyler HS, leads the highest socioeconomic school in the study, but still experienced these investment challenges:

> I think that the . . . one of the bigger challenges would be that those folks who aren't quite as invested don't feel that same level of expectation, same level of commitment that those other parents do and those other families do. And then . . . which trickles down to the kid. And then we've got to battle to get that kid to understand how important it is.

When a parent or guardian does not have certain expectations of the child, that results in additional work for the administrator. This was also a point of frustration for all administrators in the study. Floyd, the principal at Martin HS, stated, "If the parent's not motivated enough to help the kid, that . . . that's an extremely tough situation." Finding ways to get parents motivated was an arduous task for most of these administrators. They expressed that most parents were only motivated by extrinsic factors. Don expressed this during the focus group:

> Yeah, and graduation's a big motivator . . .If there's no contact for 4 years, it . . . suddenly that last week of school, when they receive a failure letter, parents are at the door. And it's like, well, where have you been?. . . . You know what I mean . . . for 4 years? Because that . . . that's a big motivator.

External motivators helped school administrators navigate parental investment barriers, but this required extra work on the part of the school. This expectation altered the school's role, from the administrator's perspective, from a purely academic institution to a more encompassing one. Marty is the director of education at Adams School District, which is located in Pierce City. Pierce City is an urban environment consistently in the top 10 of most depressed cities in Pennsylvania. On the idea of what is the school's role in parental investment barriers, he stated, "Um, let's be honest. Schools are no longer just an academic institution. We have to become a community resource." This point was reiterated by most participants, with some even going further than describing "a community resource." John is the principal at Taylor HS, which is located in the suburb of Pierce City. He faces some of the same parental investment challenges of non-nuclear families that Adams School District has, but John's frustrations can be seen with the additional expectations placed upon the school. John stated that schools have to be the child's nutritionist, chef, health care provider, and physical trainer. He went on to express that schools are asked to provide every service including being a parent. All participants experienced this demand of additional roles above academic obligations. The most rural participant in the study was Dave at Madison HS. He also experienced this expectation of the school being expected to parent.

However, this expectation of being the academic institution as well as fulfilling all other expectations can be a difficult one to meet. Shawn addressed this balancing act:

> People don't like people telling them how to do things in their own homes, you know. And I

> think that's just one of those fine lines that I . . . I think that we can consistently offer support and help and . . . again, I go back to the . . . when are we being the know-it-alls and intrusive on families' lives.

With this fine line to navigate and the additional expectations placed on the school, administrators are tasked with much more than simply running an academic institution.

EXTRA SUPPORTS AND TASKS

Some administrators in the study stated that once they know that a student is from a non-nuclear situation that has some challenges, they put supports into that child's schedule. As a result of the extra expectations placed on schools, school administrators are tasked with finding additional supports to assist that student and family. Floyd was one of those administrators who stated that when a student is identified as not having a nuclear family, he attempts to put supports into that child's schedule. Sometimes, extra support is in the form of taking time out of the administrator's day to listen to the child. Jake described his experience with supporting a child:

> But if all of those things are running in your head and the kid is staring out the window, they're thinking about all the possibilities. And so many times when . . . parents are afraid to tell their kids what's going on, so. . . . And the kids make up the worst possible case scenario you could ever imagine. They imagine the worst things ever. And you call home and you're like, "Hey, this stuff is going on in this kid's head." And, um, they're like "Ahh no, it's not that, it's this." And the kid is like, "Well, I didn't know that you . . . I didn't know that." And there's a communication breakdown there. So kinda be that communication bridge for kids. And then they're like, "Oh." Like, "You're good?" "Yeah." "Go back to math class."

Almost all of the participants in the study expressed a situation similar to the one that Jake described. Every school in the study had a social worker, which the participants described as vital to serving these non-nuclear families. Tom, the principal at Fillmore HS, talked about the social worker role in serving these families. He believes that families expect that support from the school. However schools are limited in their resources and the guidance counselor and social worker are becoming overwhelmed with the mental health issues presented to them. Tom went on the express that the issues that the support staff are assisting children with are issues that nuclear families do not have. This support role is so crucial to the school's newly discovered role of serving non-nuclear families that Tom is weighing the social worker role over the role of a teacher:

> We may have an English teacher retiring here next year. I'm actually on the fence of saying, "Let's not replace with an English teacher, but let's replace with a social worker." Because there's way more work on the plate for those guys. Like I said, we have families . . . we have classes where maybe there's maybe one nuclear family or none . . . So, I mean, it's like the need is there for us to save these kids in the classroom.

The social-emotional piece of schools is an added burden that takes away from the education that needs to be provided to meet the demands of state testing and the child's academic growth. During the focus group, Judy expressed that if she had a magic wand, this would be the piece that she would remove, "take that piece away from the schools, or support that piece in the schools, by some outside body So that we would have time to do more with other kids and those kids."

COMMUNICATION BARRIER

The child's and families social-emotional needs result in an additional service the school has to provide. However, that is not the only additional task placed upon the school. As a result of the dissolving nuclear family, multiple contact persons are involved in that child's life. Therefore, this also provides an additional challenge for the school administrator. Jake expressed that he believes that communication is the number one barrier he experiences with non-nuclear families. When that breakdown in communication takes place, it creates a barrier for the administrators to navigate. Dave shared his experience:

> So, in a lot of non-nuclear situations, I maybe have a very lengthy list of contacts because for whatever reason, somebody is the primary caregiver for this child . . . but wasn't naturally set to [be]. I'm thinking grandparents or other things. And it takes a village, so to speak, maybe to raise that child . . . Then on a nuclear situation, sometimes I have the traditional. There's a mother and a father listed. There's one workplace here and one workplace here, and it's more simple. So there's definitely a communication barrier, more or less getting to that person, um, sometimes a little bit with getting a response or a motivation to get back.

This phenomenon of multiple at-home contacts was a familiar one among all participants. Walter explained in order to address that communication barrier it is vital to establish the best way to communicate. Identifying the individual that you need to talk to in order to best support that child and whether that is by email, a phone call to the house, a phone call to a cellphone, etc. helps alleviate some of the resistance administrators face with communication barriers.

The challenge that the participants found was that contact information was not being shared. Jasmine, the principal at Monroe HS, wrestled with how to obtain this information from non-nuclear parents:

> There is kind of that fine line between asking too many personal questions and trying to make [sure] everyone's involved. Uh, because, you know, if you know a family is separated, and . . . you can't really say to Mom like, "Well, do you want Dad to be included in this?" 'cause then that's awkward.

The "awkward" conversations can sometimes lead to significant issues. A similar experience was expressed by Donna during the focus group discussion about reaching out to parents. She experienced several situations where the biological parent wanted information on the child but was not on the approved contact list because of a decision by the other biological parent. When dealing with domestic disputes such as the one Donna experienced, these situations can escalate, requiring the injunction of a court order. This requires additional effort on the administrator's part to ensure that all vested and required parties are involved with that

individual student's communications. John stated,

> Upon learning that a student comes from a non-nuclear family, I think about the potential struggles that a student may face, the possibility of legal orders that may need to be identified and/or followed, and I think about the potential data that we need to collect to ensure total family communication is enabled and followed.

Court orders add another dimension, but the school administrators must rely on parents to provide that information. There is no way to validate the information provided by the parents and the actual court rulings set forth within the current system. Walter had an experience where parents were not following the court order and had written up their own agreement. This was frustrating for Walter because he is required to follow the court order. The overall sentiment from all the participants was a desire to help the student and communicate with whomever is needed to achieve that goal. However, frustration mounted whenever information was not shared with the school.

Gathering the required information from parents, either on legal proceedings or basic contact information, can be a challenge. Furthermore, keeping that information up to date is another hurdle that required additional effort from the administrators and additional frustration when the information is obtained. Floyd shared the same sentiment of others in the study about the onus of updating information falling onto the parent when he stated, "Communication is always a challenge. And, it's a challenge if the other part of the family doesn't do their part to make sure that their information is updated or their email is updated or their address is updated."

Most of the participants understood the challenges for parents of constantly updating the information at the school. Tom explained that updating the information at the school may not be a priority when you are moving. The change in addresses could be for a multitude of reasons. "You know, families change, and the dynamics change so frequently," Don stated. Another reason for the change of address was families being transient. Floyd explained the frustration of a transient non-nuclear family as well as the added dynamic of the COVID-19 pandemic:

> The family who would move from, from this house to this house, and they do that every 2 months, but they won't tell the school. So we're sending letters home, and post office stops forwarding them after a . . . after a year or after a month, and then, then we get the, um, the letters back. And the only way that we got those address changed is if we called the kid in and said, "Where do you live?" I'm serious. And then it was a real son of a bitch during COVID 'cause I didn't have the kids to talk to. That was absolutely just horrendous.

The challenge of updating student contact information and ensuring that the proper guardian was informed was an issue for all participants, regardless of the school district. This communication barrier led to more significant issues requiring even more additional work from the administrator. The approach of addressing the communication barrier by utilizing the student even had challenges. The student may be manipulating information to their advantage, not reliable enough to relate the information, or regulating the information for their own situation. The topic of the child being relied upon for the communication was

discussed within the focus group while discussing parental contact. The group expressed that sometimes the lack of communication is purposeful from the student standpoint so that only one side of the family knows about an event or situation.

Regardless of the path to go around the barrier of communication, all of the administrators were dedicated to finding a pathway for the child to be successful and the parents to be involved. This may have resulted in extra effort and frustrations, but the end product was to help the child. Walter best summarized the barrier of communication:

> You know, education is different than it was 30 years ago, too. So, if we said we were gonna teach the same curriculum that we did 30 years ago, that would not be acceptable. So, if we say we're gonna communicate with families the same way we did 30 years ago, that's not acceptable either.

JUDICIAL SYSTEM

Another additional task placed upon administrators when helping non-nuclear families through parent investment barriers is having Student Attendance Improvement Plans (SAIPs). These meetings are required by the state of Pennsylvania for students who miss a specified number of days. Dave stated that these meetings are common among non-nuclear families, "I'm much more prone to probably have one of those SAIP meetings with a non-nuclear family for whatever reason, the expectation of making it to school or, or actually accomplishing the mission." Fulfilling the state requirement requires the administrator to meet with the student and their parents or guardians to formulate a plan to increase that child's attendance. SAIP meetings require additional time and effort on behalf of the school administrator. These required meetings were met with mixed responses. Judy had a positive experience with the meetings and found them to be a positive use of her time:

> At first, I thought, "This is just another . . . seven-page paper I gotta complete before I have to do that. I gotta meet with the parents." But I was surprised at how something as simple as that, just a contract that we signed . . . having them come in and sign this and say, "Okay. . . now, this means that your child can't miss school, and they can't do this," that did work for some of them. So, setting that expectation with attendance, that was helpful.

This positive experience was not shared among all participants. Some viewed the required meeting as another burden that takes away from the school's role. John expressed this frustration during our interview:

> Student attendance improvement plan meetings are a waste of time. I'm still the one who's doing all the work...And the worst thing that's gonna happen is the student's gonna be put on juvenile probation. They're gonna have to meet the juvenile probation officer at school. They're going to have to do things at school. . . There's no repercussions on parents. If the parents lost money or if the parents lost time away from their job because they have to do this or they have to do that or they have to do this or if the parents got fined, that is a bigger deal. That is more realistic. But as it stands now, the state's just added more and more work to the schools, more and more work to the court system.

John's mention of the court system was not the only instance of the school and judicial systems overlapping. Not only was working with another institution a challenge, but it also required additional time going to the court. The cross-over with the judicial system also requires additional knowledge about the different laws and courts about which school administrators are not trained. While talking about the SAIP meetings, the focus group began to discuss the frustrations with the court system. The group shared that a majority of the students that they have to take to court for attendance issues are from non-nuclear families. However, they expressed frustration with the court system because not all judges will treat truancy the same. With a lack of consequences from the judicial system, the school administrator is tasked with addressing an attendance issue with little to no repercussions for the child or family.

Regardless of the barrier, whether navigating communication challenges, scheduling and hosting SAIP meetings, or working with the judicial system, the school administrator is expected to put forth extra effort to assist the non-nuclear family. This extra effort and support are in accordance with path-goal theory, which suggests that the leader attempts to make the path to success clear and easy (Northouse, 2019). Although the extra work falls on the administrator's shoulders, the extra effort results in a clearer route to success for the non-nuclear family.

DISCUSSION

The purpose of this study was to better understand school administrators experiences with parental investment barriers of non-nuclear families. To this end, an analysis of school administrators' experiences gave a voice to a perspective that was rarely considered in previous research. This analysis resulted in the theme that school administrators are faced with additional tasks and efforts to be put forth that are not academic in nature.

Non-nuclear families have become more commonplace in American society (Smock & Schwartz, 2020). Coontz (2016) found that these societal changes have made it harder for parents to raise children without assistance. In this study, school administrators expressed that the assistance these families seek usually is expected to come from the school system. Israel et al. (2001) found that rarely does the government provide resources to public schools in order to help the family in non-academic matters. This was ubiquitous with the participants' experiences as many expressing their desire to help families but the lack of resources in order to do so.

IMPLICATIONS AND RECOMMENDATIONS

The concept of family is always in a state of transition (Smock & Schwartz, 2020). Therefore, school administrator preparatory programs should include a component that prepares future leaders in the field of family dynamics. Research-based instruction on how to best interact with a non-nuclear family, handle communication barriers, and address barriers to parental investment would better prepare future school administrators in a changing educational environment. Previous literature has indicated that children of divorced or separated parents have lower academic achievement than do their nuclear family counterparts (Laursen et al., 2019). If there is an identified population of disadvantaged students, training should be available to help this subgroup. Participants continued to express throughout the study that their only training was through their own experiences, whether personally or professionally. Therefore, higher education

institutions need to reconsider their school administrator training curriculums. These programs should include a component on the majority family structure in America: non-nuclear families. Training programs should also include proper and effective techniques for addressing parental investment barriers of non-nuclear families.

Based on the implications of the results of this study and the increasing number of non-nuclear families in American society, further research is recommended to understand better the challenges school administrators face with addressing the needs of this growing subgroup. The findings of this study implicated the need for social services in the school district as well as additional funding for services currently being provided. Research from the non-nuclear household perspective on the appropriate level of intervention from the school would help clarify the school's role.

DELIMITATIONS AND LIMITATIONS

The purpose of this transcendental phenomenological study was to describe the lived experiences of school administrators in Central Pennsylvania with leading non-nuclear families through parental investment barriers. Delimitations include utilizing a purposeful sample, site selection, and the scope of the study. The researcher used purposeful sampling because he sought erudite participants so that the researcher could understand the phenomenon as experienced by the individual. The purposeful sampling strategy used for this study was a key-knowledgeable strategy. Key knowledgeable sampling enables the researcher to identify trends and future directions of the phenomenon (Patton, 2015). The researcher sampled school administrators with at least 5 years of high school experience and 2 years of administrative experience so that they had experience with the phenomenon from an administrative perspective before the COVID-19 pandemic. Schools in Central Pennsylvania were selected to ensure that bias would not affect the study because of the researcher's relationships with administrators in the area.

There were also limitations in this study. The purpose of this transcendental phenomenological study was to collect in-depth descriptions of the lived experiences from a small number of participants (Creswell & Poth, 2018). Data saturation was achieved with 12 participants (Creswell & Poth, 2018). One limitation of the study was the lack of diversity: All participants in the study were White. This sample may not have been representative of all school administrator experiences; however, the sampling was representative of Central Pennsylvania administrators. All participants agreed to participate in this study voluntarily. Therefore, they may have been more confident in their ability to lead non-nuclear families through parental investment challenges than a potential participant who did not agree to be in the study. Finally, the geographical location presented a limitation to the results because the school administrators were all located in Central Pennsylvania.

CONCLUSION

Results indicated that school leaders need more support and financial assistance. Participants in this study experienced additional tasks associated with assisting non-nuclear families through parental investment barriers. Participants attempted to negate the parental investment barriers for non-nuclear families through barrier removal and support services. School administrators are on the front line of the battle for students from non-nuclear families. Tom summarized that participants viewpoint with "the end goal is to assist the student, so they can be as successful as they wanna be or can be. I mean, that's your end goal, is to help the kid. So whatever it takes then to engage families." What it takes is these everyday heroes having additional training and support.

REFERENCES

An, B. P., & Sorensen, K. N. (2017). Family structure changes during high school and college selectivity. *Research in Higher Education,* 58(7), 695–722. https://doi.org/10.1007/s11162-016-9445-8

Brock, S., & Edmunds, A. L. (2010). Parental involvement: Barriers and opportunities. *EAF Journal*, 21(1), 48–59.

Bengtson, V. L. (2001). Beyond the nuclear family: The increasing importance of multigenerational bonds. *Journal of Marriage and Family*, 63(1), 1–16. https://doi.org/10.1111/j.1741-3737.2001.00001.x

Cavanagh, S. E., & Fomby, P. (2019). Family instability in the lives of American children. *Annual Review of Sociology*, 45(1), 493–513. doi:10.1146/annurev-soc-073018-022633

Coontz, S. (2016). *The way we never were: American families and the nostalgia trap.* Basic Books.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). Sage Publications.

Garbacz, S. A., Herman, K. C., Thompson, A. M., & Reinke, W. M. (2017). Family engagement in education and intervention: Implementation and evaluation to maximize family, school, and student outcomes. *Journal of School Psychology*, 62, 1–10. https://doi.org/10.1016/j.jsp.2017.04.002

Gordon, K., & Fefer, S. (2019). Discipline history and demographics: Which factors relate to school climate perceptions among high school students? *School Psychology Forum, Research in Practice*, 13(1), 16–28.

Hamlin, D., & Flessa, J. (2018). Parental involvement initiatives: An analysis. *Educational Policy*, 32(5), 697–727. https://doi.org/10.1177/0895904816673739

House, R. J., & Mitchell, T. R. (1975). *Path-goal theory of leadership*. https://apps.dtic.mil/dtic/tr/fulltext/u2/a009513.pdf

Israel, G. D., Beaulieu, L. J., & Hartless, G. (2001). The influence of family and community social capital on educational achievement. *Rural Sociology*, 66(1), 43–68. https://doi.org/10.1111/J.1549-0831.2001.TB00054.X

Laursen, L. L., Madsen, K. B., Obel, C., & Hohwu, L. (2019). Family dissolution and children's social well-being at school: A historic cohort study. *BMC Pediatrics*, 19(1), 449–458. https://doi.org/10.1186/s12887-019-1821-z

Lim, E., Davis, J., Choi, S. Y., & Chen, J. J. (2019). Effect of sociodemographics, health-related problems, and family structure on chronic absenteeism among children. *Journal of School Health*, 89(4), 308–318. https://dx.doi.org/10.1111%2Fjosh.12736

Lunenburg, F. C. (2011). Expectancy theory of motivation: Motivating by altering expectations. *International Journal of Management, Business, and Administration*, 15(1), 1-4.

Miller, G. H. (2020). Future trends in marriage and family. In D. N. Farris & A. J. J. Bourque (Eds.), *International handbook on the demography of marriage and the family* (pp. 307–315). Springer Publishing.

Moehling, C. M. (2004). Family structure, school attendance, and child labor in the American south in 1900 and 1910. *Explorations in Economic History*, 41(1), 73–100. http://dx.doi.org/10.1016%2Fj.eeh.2003.07.001

Myers, S. M., & Myers, C. B. (2015). Family structure and school-based parental involvement: A family resource perspective. *Journal of Family and Economic Issues*, 36(1), 114–131. https://doi.org/10.1007/s10834-014-9409-0

Northouse, P. G. (2019). *Leadership: Theory and practice* (8th ed.). SAGE Publications, Inc.

Olowoselu, A., Mohamad, M. A. B., & Aboudahr, S. M. F. M. (2019). Path-goal theory and the application in educational management and leadership. *Education Quarterly Reviews*, 2(2), 448-455. http://dx.doi.org/10.31014/aior.1993.02.02.77

Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Sage Publishing, Inc.

Razik, T. A., & Swanson, A. D. (2010). *Fundamental concepts of educational leadership and management* (3rd ed.). Allyn & Bacon.

Santin, D., & Sicilia, G. (2016). Does family structure affect children's academic outcomes? Evidence for Spain. *The Social Science Journal*, 53(4), 555–557. https://doi.org/10.1016/j.soscij.2016.04.001

Smock, P. J., & Schwartz, C. R. (2020). The demography of families: A review of patterns and change. *Journal of Marriage and Family*, 82(1), 9–34. https://doi.org/10.1111/jomf.12612

Sun, Y., & Li, Y. (2011). Effects of family structure type and stability on children's academic performance trajectories. *Journal of Marriage and Family*, 73(3), 541–556. https://doi.org/10.1111/j.1741-3737.2011.00825.x

Vroom, V. H. (1964). *Work and motivation*. Wiley.

ABOUT THE AUTHOR

Brian Randall, Ed.D., serves as a high school instructor at Portage Area School District and an adjunct professor at Penn State University – Altoona. He has over twelve years of experience in public education. Dr. Randall completed his doctorate in the Fall of 2021. He has obtained his principal certification and is actively pursuing an administrative position. Dr. Randall is passionate about engaging families and creating learning environments that foster engagement.

brandall3@liberty.edu

IT'S ALL ABOUT TO CHANGE:
IMPLICATIONS OF REFORMING GRADING AND ASSESSMENT

Divonna M. Stebick, Ph.D.
Megan L. Pilarcik, M.S.
Daniel W. Hartman, Ed.D.

Pennsylvania PASCD
LEARN • TEACH • LEAD

PENNSYLVANIA EDUCATIONAL LEADERSHIP | VOLUME 41, NUMBER 1

ABSTRACT

Calls to reform grading systems and other assessment practices have been growing for several decades. There is consensus among many educators that grading and assessment practices that have been traditionally accepted as good practice are at best ineffective and at worst have a negative impact on raising achievement. Consequently, there is no single solution or methodology for grading that has emerged as the best practice. A variety of contemporary grading approaches have gained widespread popularity in recent years, typically being referred to as standards-based grading, standards-referenced grading, proficiency-based grading, or competency-based learning. A challenge, however, is that different school districts define these terms differently and models vary widely in their implementation. Although there is no single grading model that is the panacea for all of the ills of past practices, the researchers identified the common ground about what an effective system for grading and assessment should include. Even though further study is needed, the resounding evidence indicates that contemporary grading practices such as standards-based grading are a preferred model if the goal of grading is to accurately communicate student learning and achievement.

INTRODUCTION

When I started teaching thirteen years ago, my classroom at South Western looked like any other typical classroom: I graded homework, thought tests should represent the majority grade, and penalized late and missing work with zeros. Students needed to be taught responsibility and using their grades as a reward/punishment system was an impactful way to do just that. This exact reason was why I graded homework as well—if there was a grade, my students were more likely to complete it.

~Thoughts from high school teacher at onset of assessment and grading reform

Grading reform has been a common, and yet one of the more controversial, topics of conversation in K-12 education for the past two decades. A plethora of literature and research has been produced articulating the inaccuracies of traditional grading practices for assessing student learning and the need for changes on how to measure and report student performance.

Many researchers and experts emphasize the need to change the methods in which schools measure student grading through traditionally accepted grading scales and calculations towards methods that more accurately measure students' current level of mastery of specific standards or learning targets that rely less on averaging scores across a period of time and that eliminate one overall omnibus course grade that is an average of scores for all assignments for all content during a specified period of time (Guskey, 2013; Marzano, 2000; Marzano & Heflebower, 2011; Miller, 2013; O'Connor & Wormeli, 2011; Reeves, 2011). Other works advocate for eliminating grading practices considered to have negative impacts on learning outcomes, or that are considered toxic, such as giving "zeros" for late or missing assignments, not allowing students to retake tests or re-do assignments, giving extra credit for tasks not directly aligned to intended learning outcomes, giving points for non-academic factors such as behavior or motivation, and grading assignments and tasks that are more formative in nature and meant to allow students to practice and develop their skills (Fisher et al., 2011; O'Conner & Wormeli, 2011; Reeves, 2004; Schoen et al., 2003; Townsely & Buckmiller, 2016; Wormeli, 2011). Finally, others point to a need for improved assessment literacy among educators and improved assessment design, including an increase in more authentic, performance-based, assessments (Guskey, 2005; McTighe, 2018; Stiggins, 2004). Assessment literacy is not only necessary for teachers but for all stakeholders (Davison, 2013). The purpose of assessment must be to clearly connect assessment literacy with key responsibilities of each stakeholder (Popham, 2009). Stakeholders have different needs which demand different levels of assessment literacy, therefore, the most critical factor contributing to failure of assessment reform is the misconceptions (Davison, 2013). These misconceptions lead to poor implementation as simply changing the assessment practices of teachers will reform the assessment process, whereas, the district need to establish a strong assessment reform culture with all stakeholders across all levels of the district (Davison & Leung, 2009).

Although there is consensus among many educators that grading and assessment practices that have been traditionally accepted as good practice are at best ineffective and at worst have a negative impact on raising achievement, there is no single solution or methodology for grading that has emerged as the best practice. A variety of contemporary grading approaches have gained widespread popularity in recent years, including standards-based grading, standards-referenced grading, proficiency-based grading, and competency-based

learning. A challenge, however, is that various organizations define these terms differently and models that are implemented in school organizations vary widely. Although there is no single grading model that is the panacea for all of the ills of past practices, there is some common ground when it comes to beliefs about what an effective system for grading and assessment should look like, including the beliefs that (Pollio & Hochbein, 2015; Peters & Buckmiller, 2014):

- Assessment and grading should enhance learning, not just measure it. Teachers, students, and others should be able to utilize assessment results to identify learning misconceptions and needs and inform next steps for instruction and learning.
- The primary goal of grading assessments is to communicate student achievement and learning. Grades should be an accurate portrayal of a student's current level of their mastery towards identified desired learning outcomes.
- In order to accurately reflect what students know, understand, and are able to do, grades should only reflect progress towards standards and learning outcomes, and not reflect non-academic factors such as attitude, effort, and behavior.
- Students should be involved in the assessment process through means such as having choice in assessment methods, self-evaluation and self-reflection, and through monitoring and tracking their own progress towards a learning goal.

Although calls to reform grading systems and other assessment practices have been growing for several decades, the field lacked empirical studies into the impact contemporary approaches to grading, such as standards-based systems, had on achievement and learning outcomes. Most recently, however, more studies are being conducted on grading practices providing evidence on the effectiveness of contemporary grading models, as well as the characteristics and practices within these systems that impact their level of effectiveness.

Standards-Based Grading; Evidence of Effectiveness

Research is still minimal on the direct impact contemporary grading practices such as standards-based grading has on student achievement, and evidence is emerging that suggests such approaches do provide a more accurate representation of student achievement and performance. Pollio and Hochbein (2015) note that most of the available research on grading describes teachers' perceptions of grading practice absent a rigorous examination of the influences of classroom grades on learning.

In a recent study of high-needs high schools in Kentucky which implemented standards-based grading practices, stronger correlations between standardized test scores and course grades were more evident than they were prior to switching to a standards-based system (Hochbein & Pollio, 2016). An earlier, unrelated, study also conducted in Kentucky, Guskey et al. (2011) found that teachers in schools where standards-based report cards were implemented overwhelmingly agreed that a standards-based approach provided better and clearer information about students' learning. This same study found that parents and families also favored the standards-based approach over the traditional model.

Because of the fact that standards-based approaches to grading often do not rely on a mean score of student work, rather implement a formula for median or mode to calculate a final grade, they are often considered more subjective than traditional grading practices (Pollio & Hochbein 2015). Supporters of

contemporary grading practices dismiss this as a false reality, pointing to what they consider toxic practices that are commonplace in traditional grading practices that lead to more subjectivity, such as allowing some students to complete extra-credit, giving points for behavior and effort, etc. O'Conner (2011) concedes that because grading involves human judgment, there will always be subjectivity in grading, and that even calculated grades are subjective ratings that give the appearance of an objective measurement. Recent studies have shown that teachers' informed professional judgments yield a higher reliability and more consistency in determining grades than computer-based programs that produce scores based on statistical algorithms (Guskey & Jung, 2016). Subjectivity is expected in grading so long as the grading decisions can be defensible, meaning that they are supported by credible evidence.

Although the impact these grading systems have on their own on learning and achievement may not be clear, having measures that are a more accurate measure of learning can help to support practices and strategies in which the research shows a clear positive impact. In a meta-analytic synthesis of studies conducted on the impact of various instructional strategies, tracking student progress through the use of scoring scales was found to have the most significant impact on achievement, resulting in a 31% gain (Haystead & Marzano, 2009). Effectively tracking student progress requires accurate measurements of students' current level of learning which the research has indicated standards-based approaches are better equipped to provide than traditional grading models.

Although further study is needed, the resounding evidence is indicating that contemporary grading practices such as standards-based grading are a preferred model if the goal of grading is to accurately communicate student learning and achievement. The pursuit of an accurate system to measure learning has caused many districts and schools to investigate, develop, and implement new grading and assessment practices. The Mustang Coral School District, was one of these districts, which has resulted in a multi-year process of reimagining and redesigning their district practices and policies in regards to grading and assessment. The authors of this study worked on the district's grading and assessment committee and included a district administrator, a high school mathematics teacher, and a parent.

Redesigning a Grading System; One District's Process

The grading and assessment change initiatives began as a grass roots effort within the Mustang Coral School District. In actuality, the genesis of these efforts originated with a group of educators getting together for a different purpose all together, unwittingly laying the ground work for what would lead to district-wide grading and assessment reform. During the 2007-2008 school year, a group of high school teachers and administrators began meeting to discuss best practices the topic of mid-term and final exams. This group was delving into questions such as "If students demonstrate mastery throughout a course, should they be required to take a final exam?, How much should a final exam be weighted, or should it be weighted at all?" as well as other related questions. As this group grappled with these questions and shared information with district leaders, it became apparent that there was a greater need than simply addressing mid-term and final exam expectations, and that the district needed to take a deeper and wider look into all of its practices involving assessment and measurement of student learning.

This identified need gave rise to the formation of a district-wide "Grading and Assessment Committee" beginning in the 2008-2009 school year. Over the next several years, this group engaged in exploring effective grading and assessment practices through participating in book studies, reviewing research reports and articles, and conducting site visits to districts that were implementing standards-based grading systems or other contemporary grading practices. The committee then came to consensus on the following list of belief statements about grading and assessment that were adopted by the district:

1. The purpose of assessment is to improve learning.
2. Teachers will use assessment and grading practices that accurately measure student performance.
3. Grades will reflect what students know, understand, and are able to do.
4. Grades will be based on a variety of high-quality common summative assessments.
5. Students should be involved in the assessment and grading process.

In order to bridge these beliefs into classroom and instructional practice, the district also developed the following comprehensive list of best practices which served as non-negotiable expectations:

Developing and Utilizing Assessments

- Align assessments to clearly defined learning targets.
- Base learning targets on state or national standards and anchors.
- Use frequent and varied, collaboratively developed common summative assessments to measure students' knowledge and skills accurately.
- Provide multiple opportunities for students to show understanding, after re-teaching and/or remediation.
- Continually review and revise all assessments for curriculum alignment and quality.

Using Assessment Results

- Collaboratively review assessment results to plan and prioritize future instruction.
- Use feedback to move students forward in their understanding.
- Communicate to students what was done well and what should be done to improve.
- Give feedback on both formative and summative assessments.
- Use assessment results to prioritize future differentiated instruction.
- Use evidence from summative assessments to determine course grades.

Reporting Academic Achievement

- Report behavior, effort, and participation separately from academic achievement.
- Base grades on summative assessments and tasks.
- Base report card grades on the most current evidence based on the attainment of learning target.
- After reassessing, record the most current summative assessment score as the grade.

Responding to Late Assignments

- Score late work for achievement, not punctuality, which should be reported separately.
- Record incomplete work as "Incomplete" until it is made up.
- Use available resources to get students to complete missing/late work.
- Communicate with parents when students are not completing work.
- Set a specific deadline for incomplete work to be made up; until the work is completed, use appropriate consequences to encourage work completion and to change future behavior.
- If the work is not made up by the set deadline, record the grade based on existing evidence of the student meeting the learning goal. Lack of evidence may result in no credit earned.
- In extreme cases, teachers should refer students to the student concerns process.

Utilizing Homework as Formative Assessment

- Use homework as practice, previewing, or extension of material and/or skills, with the purpose of providing information/feedback to students.
- Base the amount and type of homework on the age and needs of the students.

Involving Students in the Assessment and Grading Process

- Students should be expected to take ownership of the process through tracking their own progress, goal setting, self-assessment, and requesting opportunities to be retaught and retested.
- Students should be taught to use formative assessment results to improve and adjust their approaches to learning.

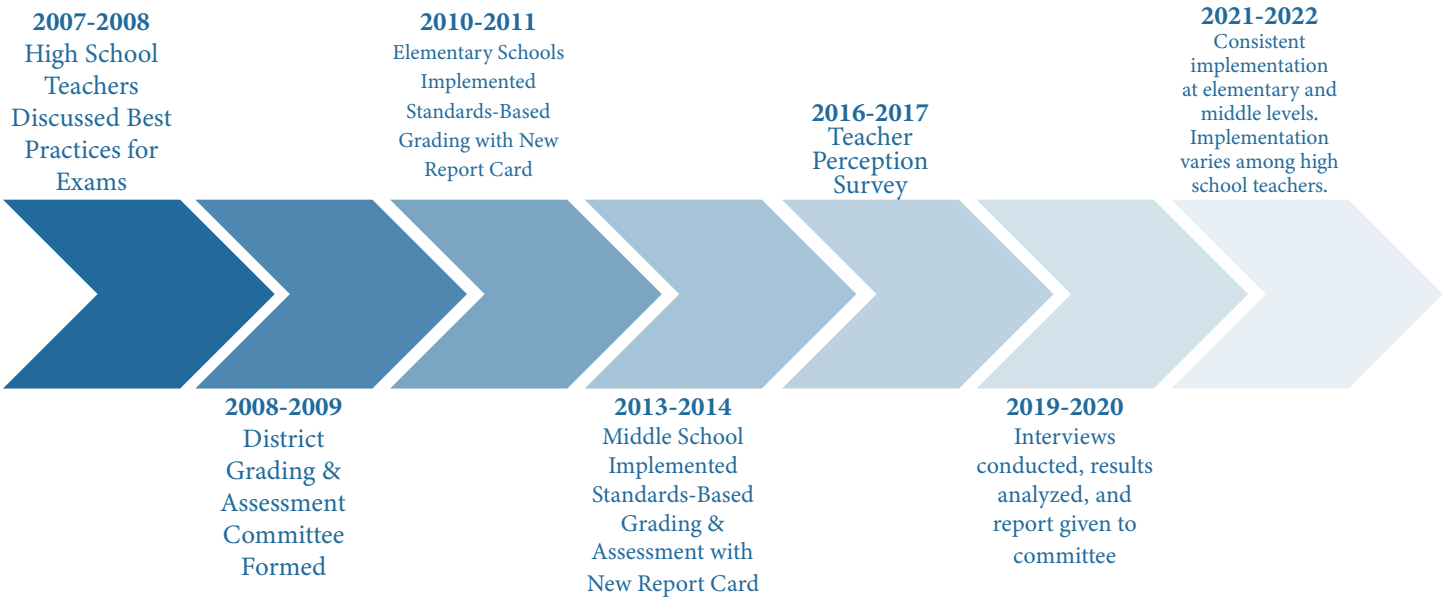
The district believed that in order for teachers to be able to successfully transition their practices to align with these new expectations, there would need to be substantial professional development and support. Peters and Buckmiller (2014) reported that successful implementation of any grading reform should be purposeful and well communicated, and include a high level of support including professional development and collaboration. Professional development efforts throughout the district focused primarily on grading and assessment during these initial years of implementation.

In addition to adopting these identified best practices, beginning in the 2012-2013 school year the district elementary schools (K-5) implemented a standards-based grading system and report card. The district's intermediate school (6-8) followed suit the next year, in 2013-2014. The key features of this new grading system at both levels were that (a) students no longer received a single course grade, but a grade for different standards categories identified for each grade level and course based on clusters of standards; (b) for each reporting category, students would be graded on a 4-point scale (see appendix) instead of the traditional 100-point scale A, B, C, D, F model; (c) final grades for each category were to be based on students' current level of mastery determined by the teacher's analysis of multiple pieces of evidence, no longer an average of scores from a designated reporting period; and (d) in addition to standards-based grades based on academic conduct, students received a separate 'Habits of Learning' grades which measured factors such as motivation, participation, and communication.

As is with many district grading reform efforts, challenges at the high school emerged and eliminated the possibility of an effective implementation. These factors included trust, early involvement, due diligence (Foos et al, 2006), personal interest, shared values (Dhanaraj et al., 2004), intrinsic motivation (Osterloh & Frey, 2000) and a fit in the Mustang Coral School District's secondary curriculum (Kippers et al., 2018). Reflecting on the lack of change at the high school, it was clear that the faculty did not identify and adopt the philosophical changes that would move the assessment practices to a standards-based-grading process. Therefore, the high school became the place where standards-based-grading died. The faculty prefers the traditional 100-point 'A, B, C, D, F' grading scale. Several factors, including parent and community support, and perceived barriers such as a negative impact on college admissions and scholarship opportunity have presented challenges that have contributed to the high school from making such a transition.

During the 2016-2017 school year, the district's grading and assessment committee determined that there was a need to gage teacher perceptions on the grading in order to identify professional development needs related to assessment literacy and effective grading practices, and to identify potential revisions to district grading policies and practices. In order to elicit the most valid and reliable feedback possible, district administration and the grading and assessment committee partnered with the principal investigator, the parent, to develop a qualitative study that focused on teacher perceptions of the district's grading and assessment practices. A summary of the work can be seen in Figure 1.

FIGURE 1
Timeline for Study



CURRENT STUDY

Throughout the past eight years, Mustang Coral School District created an active, dynamic Grading and Assessment Committee. This committee includes faculty from all levels (elementary, middle, and high schools), administrators, board members, and parents; meeting monthly to reconstruct the grading practices, report cards, and assessment tools. Recently, it has become apparent that not all stakeholders embrace the same understanding of assessment literacy. This lack of a common understanding resulted in obstacles that include misguided grading practices, differences in the reporting of grades, and inconsistent assessment practices. The committee identified a need to complete a gap analysis of assessment literacy, so that the committee could move forward to ensure student learning and growth. The faculty responses will help the committee to capture the quality of process/product within the area assessment literacy so that the committee can support all stakeholders by providing a consistent definition, proper professional development, and clear communication.

METHODS

District Demographics

The Mustang Coral School District serves approximately 4,500 students in grades K-12, and is comprised of four elementary schools (K-5), one intermediate school (6-8), and one high school (9-12) in a rural area. The student population is made up of 87.7% students identifying as Caucasian, 4.7% Hispanic, 2.5% African-American, and 1.8% Asian. Approximately 32% of district students come from economically disadvantaged families, with individual schools ranging from 24.3% economically disadvantaged to as high as 45.5%. The district's graduation rate is 96.2%, with approximately 78% of students going on to post-secondary education upon graduation.

The district teaching staff is comprised of 309 full-time faculty members. South Western's faculty includes individuals with a wide range of experience, with 16% of faculty having less than five years of experience, and 23.1% having 20 or more years of experience. Across the district, 82.7% of faculty members hold a masters degree or higher.

PARTICIPANTS

The committee solicited interest for participation in the study through an interest survey. Based on the number of responses from each school and the amount of time the interviewer had to conduct individual interviews, the researcher used a stratified random sampling process to invite 100 teachers to participate. After multiple invitations, 81 teachers agreed to participate in the interviews. Each interview was conducted by the researcher at one of the four elementary schools, middle school and high school. Each building administrator secured a quiet rooom where the interview could be conducted with minimal distractions. The interviews lasted on average twenty minutes, ranging from 15 minutes through 55 minutes. In order to ensure that teachers felt safe and non-evaluated; the principal investigator conducted each interview, she is not employed by the district but represents parents on the committee. The interviews were audio-recorded then transcribed. Across the district there were 293 teachers who administer student grades and therefore were eligible to participate in the study. To ensure fair representation of teachers from grades and content areas, the researcher used a stratified random sampling process. The researcher sent a total of 146 invitations to teachers in the

district. Of these 146 invitations, 26 warranted no response, 42 teachers declined to participate leaving 81 willing to participate. The participants included 39 elementary teachers, 21 middle school teachers and 21 high school teachers. There were 60 female and 21 male teachers. Their years of teaching experience ranged from one year to 33 years. Their degrees ranged from a bachelors to multiple masters degrees.

Process

Once teachers agreed to participate, by signing an informed consent form, each participant scheduled an interview using an online scheduling tool indicating the researcher's availability. Interviews were recorded, transcribed and then read by each researcher and coded to identify common themes. The researchers discussed their individual coding to identify common themes of relevance. Transcriptions were reread by each researcher and coded with identified themes to ensure inter-rater reliability. The identified themes, gathered across all interviews, were organized into two categories; assets and challenges.

FINDINGS

The identified assets included: *meeting student needs, specific feedback*, and *student reflection / goal-setting.* The first asset, *this assessment process meets student needs,* was identified by 100% of the elementary teachers, 92% of the middle school teachers, and 81% of high school teachers. The most common comments are captured below:

- *Look at individual child instead of the whole class, know my students more with less grading, spend more time focusing on each child.*
- *Challenge my belief that grading is more about learning and not behavior. Behavior is subjective.*
- *Can meet individual student needs.*
- *Identifies students' growth.*
- *Mastery perspective is more aligned with customization, but report card does not reflect this.*
- *I now know exactly what students do well.*
- *Can provide more personalized learning.*

The second asset, *specific feedback*, can be defined as the opportunity to provide specific feedback identifying student strengths and areas of concern. This asset was identified by 85% of the elementary teachers, 67% of the middle school teachers, and 67% of high school teachers. The most common comments are captured below:

- *[Student] obstacles are managed through conversations.*
- *Separating behaviors from academics was at first an obstacle, now a benefit.*
- *Narrative feedback I provide on assessments should be transferred to report card.*
- *Would like to add space on report cards for comments in all subjects.*
- *I provide more specific feedback for students.*
- *It's about the ability now and not the grade.*
- *This model pushes teachers to give specific, constructive feedback.*

The third asset, *student reflections and goal setting*, included the opportunity for students to review their assessments to identify what they knew and what they did not know. Once the students complete their review, they reflect and then write their goals. This asset was identified by 56% of the elementary teachers, 14% of the middle school teachers, and 29% of high school teachers. The most common comments are captured below:

- *Trying to get students to look at what they learned instead of what they earned.*
- *Better [student] reflection of what was learned, since students see the behavior separated.*
- *Students write goals, work toward goals, and achieve goals.*
- *Goal writing is working!*
- *Students are self-motivated to work on goals.*
- *Cover sheets help students reflect and then set goals.*

The first challenge across all grades and subjects was the time involved to transition into the new grading and assessment process. This challenge was identified by 67% of the elementary teachers, 76% of the middle school teachers, and 71% of high school teachers. The most common comments are captured below:

- *Spent much personal time redeveloping tests and how projects are graded and assigned.*
- *Remediation time is a struggle - need intervention periods.*
- *Where do I find the time to re-assess/reteach?*
- *Turn around time to re-teach and reassess is too short.*
- *Time intensive to create and complete the cover sheets.*
- *I sacrifice instructional time to reteach.*
- *This is definitely more work.*

The second challenge identified was the drop in student motivation. This challenge was identified by 45% of the elementary teachers, 95% of the middle school teachers, and 81% of high school teachers. This challenge was most significantly shared among teachers in grades 4 through 12. The most common comments are captured below:

- *Big dip in student motivation, see that a 2 is "okay grade", parents have checked out and don't put as much pressure on children.*
- *Students are not motivated to take on this work.*
- *Students are not motivated, students don't experience failure, where is the accountability?*
- *Motivation is less now – "I'm a 2, I'm proficient."*
- *Natural consequences are a great teacher, we are stripping our children of these. This removes the urgency for learning and motivation. Developing a failure model for work ethics and removes parental responsibility.*

In the end, the biggest challenge faced by teachers across all grades was the level of inconsistency from what a "2" looks like and the implementation expectations. This challenge was identified by 45% of the elementary teachers, 95% of the middle school teachers, and 81% of high school teachers. The most common

comments are captured below:

- *Consistency of what a "2" is, is very subjective.*
- *What does the grade really mean since there is no consistency?*
- *Consistency is a struggle; narrative feedback is missing - should be added to report card.*
- *Inconsistencies in the reteach, reassess practices.*
- *Inconsistency accountability with administration.*
- *What does a "2" look like across departments, classes, grades, schools?*

The researchers sifted through the data further to identify any overlap between the greatest challenge and the greatest asset. Seventy-two of the eighty-one teachers shared that even though there is inconsistency in the process this process is meeting the students' needs as evident in student learning.

DISCUSSION

The researchers first shared the data with the district administration team, then the grading and assessment committee and most recently the board of education. During each presentation, the stakeholders agreed that time and professional development is necessary to move the process in the right direction. Knowing that this process meets the students' needs and in turn accelerates learning through customization, the district will support the teachers through professional development opportunities as well as create a clear statement to share with parents knowing that the process is asking teachers to make a practical and philosophical shift.

In practice, this includes that the district allows retests, as long as students have fully prepared for the first test but simply did not understand the concepts yet. As one teacher tells her students, retesting is not a "get out of jail free" card. Students must study and put forth their best effort on the test, but if they do not yet understand the material, they can spend more time outside of class talking, explaining, and eventually retesting. Also, teachers do not penalize late work with grades or give zeros for missing assignments without displaying due diligence in trying to obtain these assignments through assigning of intervention periods and parent contact. Teachers only grade summative assignments but frequently use formative assignments to assess learning which means we do not grade homework. Teachers do not believe in extra credit nor do they believe in behavior-based grades.

This was not an easy, overnight process of change. It involved years of conversations within the committee, schools and across the district that were grounded in research. Often, these conversations ensued conflict as expectations and implementation varied across the district. There have been, and still are, heated conversations about what a specific belief looks like in practice and if a specific structure or behavior is actually best for the students. But, from these challenging and frustrating conversations have bloomed effective practices and ideas. The district created a senior level English class, *Survival of the Fittest*, that allows students to choose a main theme, such as surviving an earthquake or a zombie apocalypse, and create a blog informing strategy, sharing first-had survivor accounts, and describing a how-to survive process. They developed a new social studies class, *Social Problems and Issues,* in which students now choose an important social issue in our area, talk to experts, and explain and propose solutions to these ideas. Their academic statistics class is structured as *Project Based Learning,* focusing on student-driven real world data analysis and issues. There

are upper level art classes that are gradeless and the focus is on the learning and identity of the created pieces instead of the score that is earned. As a result, there is an increase in students taking honors and AP level courses and more students are willing to take risks as they think in global terms.

CONCLUSIONS

It is clear that there is work to be done so that the district's assessment and grading approach is consistently implemented across all buildings, grades, and subjects. Many faculty, mostly secondary faculty, need to move from the basic questions; "What is the difference between a 93% and a 94%? What is the difference between an 89% and a 90%, or more importantly, what is the difference between a B and an A?" Teachers across all buildings continue discuss consistency and sustainability. Also, as teacher turnover becomes a more apparent challenge, the district will need to identify how to educate new staff that has not been through this process during the past ten years and how to hold the reluctant faculty accountable to implement the effective changes in all classes. The district will also need to support the faculty who have already transformed their assessment practice and adapted their course content to add relevance, whether this is a change in approach to project-based learning, a better incorporation of performance tasks, or allowing students more voice and/or choice in the products or topics they research and produce; these teachers need support.

The few faculty who have successfully transformed their assessments to align with a standards-based grading system agree that a student's grade should reflect his or her understanding of the material and not his or her behavior and that it does not matter when a student learns as long as he or she learns the material. These faculty focus on learning, rather than grading. It is evident that embracing a growth mindset ensures that teachers believe that all students can learn with the right amount of time and scaffolding. By sharing the results of this study, the researchers hope that tough conversations between faculty will challenge one another's practical and philosophical approaches to learning and assessment, and lead to a more productive, district-aligned system of assessment and grading so that student learning is accelerated.

> *Now, my classroom is incredibly different. In my academic statistics class there are no more tests. Project-based learning, with larger open-ended projects designed to meet specific learning targets have replaced tests. Students no longer need to memorize rules and formulas but rather, using their notes, need to show me that they can apply these concepts to unique and highly intriguing data sets of their own choosing. Quizzes, assessing specific content concepts, are open-note because again, the focus is on the application of ideas and not the memorization of statements. At the end of each project students conference with me, and with a previously self-evaluated rubric, they explain to me where they think they scored in each of four categories: application, presentation, communication, and reflection. We discuss these scores together before I determine a final grade with their input. If the students do not complete an assignment, it is not a zero but rather it goes in the gradebook as a "dnc" (did not complete yet) and students are assigned intervention periods, parents are contacted, and the work is eventually completed.*
>
> *~Thoughts from high school teacher after reforming her assessment and grading*

REFERENCES

Davison, C. (2013). Innovation in assessment: Common misconceptions and problems. In K. Hyland & L. Wong (Eds.), *Innovation and change in English language education* (pp. 263-275). Routledge.

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43(3), 393–415.

Dhanaraj, C., Lyles, M. A., Steensma, H. K., & Tihanyi, L. (2004). Managing tacit knowledge and explicit knowledge transfer in IJVs: The role of relational embeddedness and the impact on performance. *Journal of International Business Studies*, 35(5), 428–442.

Fisher, D., Frey, N., & Pumpian, I. (2011). No penalties for practice. *Educational Leadership*, 69(3), 46-51.

Foos, T., Schum, G., & Rothenberg, S. (2006). Tacit knowledge transfer and the knowledge disconnect. *Journal of Knowledge and Management,* 10(1), 6–18.

Guskey, T. R. (2005). Mapping the road to proficiency. *Educational Leadership*, 63(3), 32-38.

Guskey, T. R., Swan, G. M., & Jung, L. A. (2011). Grades that mean something: Kentucky develops standards-based report cards. *Kappan,* 93(2), 52-57.

Guskey, T. R. (2013). The case against percentage grades. *Educational Leadership*, 71(1), 58-72.

Guskey, T. L. & Jung, L. A. (2016). Grading: Why you should trust your judgement. *Educational Leadership,* 73(7), 50-54.

Haystead, M. W., & Marzano, R. J. (2009). *Meta-analytic synthesis of studies conducted at Marzano Research Laboratory on instructional strategies.* Marzano Research Laboratory.

Hochbein, C., & Pollio, M. (2016). Making grades more meaningful. *Kappan*, 9 (3), 49-54.

Kippers, W. B., Wolterinck, C. H. D., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, 75, 199–213.

Marzano, R. J. (2000). *Transforming classroom grading.* Association for Supervision and Curriculum Development.

Marzano, R. J., & Heflebower, T. (2011). Grades that show what students know. *Educational Leadership*, 69(3), 34-39.

McTighe, J. (2018). Three keys questions on measuring learning. *Educational Leadership*, 75(5), 14-19.

Miller, J. J. (2013). A better grading system: Standards-based, student-centered assessment. *English Journal*, 103(1), 111-118.

O'Connor, K., & Wormeli, R. (2011). Reporting student learning. *Educational Leadership,* 69(3), 40-44.

Osterloh, M., & Frey, B. S. (2000). Motivation, knowledge transfer, and organizational forms. *Organizational Science,* 11, 538–550.

Peters, R., & Buckmiller, T. (2014). Our grades were broken: Overcoming the barriers to implementing standards-based grading. *Journal of Educational Leadership in Action,* 2(2).

Pollio, M. & Hochbein, C. (2015). The association between standards-based grading and standardized test scores in a high school reform model. *Teachers College Record*, 117(11), 1-28.

Popham, W. J. (2009). Assessment literacy for teachers: faddish or fundamental? *Theory into Practice*, 48(1), 4–11.

Reeves, D. B. (2004). The case against the zero. *Phi Delta Kappan,* 86(4), 324-325.

Reeves, D. B. (2011). *Elements of grading.* Solution Tree Press.

Schoen, H. L., Cebulla, K. J., Finn, K. F., & Fi, C. (2003). Teacher variables that relate to student achievement when using standards-based curriculum. *Journal for Research in Mathematics Education*, 34(3), 228-259.

Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86(1), 22-27.

Townsely, M., & Buckmiller, T. (2016, January). *What does the research say about standards-based grading?* [Web log post]. Retrieved from http://mctownsley.net/standards-based-grading-research/

Wormeli, R. (2011). Redos and retakes done right. *Educational Leadership*, 69, 22-26.

APPENDIX

ACADEMIC GRADING SCALE
GRADES K-5

3	Your child demonstrates a thorough understanding of this skill as presented this trimester. He/she is able to perform this skill consistently and independently.
2	Your child demonstrates an understanding of this skill as presented this trimester. He/she needs assistance and/or practice to perform this skill consistently and independently.
1	Your child is in the process of building his or her understanding of this skill. He/she is not yet able to perform this skill consistently and independently.

ACADEMIC GRADING SCALE
GRADES 6-8

3	Student demonstrates a complete understanding of the skills presented in this reporting period. He/she is able to perform these skills independently, and may be partially successful at making inferences and applications that exceed expectations.
2	Student demonstrates an understanding of the skills presented in this reporting period. He/she needs additional support and practice to perform independently.
1	Student is in the process of building his or her understanding of the skills presented in this reporting period. He/she is not yet ready to perform independently

ABOUT THE AUTHOR

Divonna M. Stebick, Ph.D. is currently an Associate Professor at Gettysburg College, Gettysburg, Pennsylvania. She earned her Bachelor's Degree in elementary education at Indiana University, Indiana, Pennsylvania, her Master's Degree in Reading at McDaniel College, Westminster, MD, and her doctorate in Special Education and Literacy Policy at Union Institute and University in Cincinnati, Ohio. She taught numerous grade levels in Pittsburgh, PA, and in Carroll County, MD. Divonna researches the implications of digital literacy, grading reform, and teacher inquiry.

Megan L. Pilarcik, M.S. is currently a high school mathematics teacher at South Western High School, Hanover, Pennsylvania. She earned her Bachelor's Degree in mathematics at Gettysburg College, Gettysburg, Pennsylvania with minors in secondary education and philosophy. She earned her Master's Degree in Curriculum and Instruction at McDaniel College, Westminister, MD. Megan has taught multiple math courses but for the past many years has specialized in Statistics and AP Statistics. Megan concentrates on grading and assessment best practices, especially within the mathematics classroom.

Daniel W. Hartman, Ed.D., has served as the Assistant Superintendent of the South Western School District in York County, PA, since 2016. He has also held the positions of Middle School Principal and Assistant Middle School Principal with South Western. Dr. Hartman earned his doctorate in Educational Leadership from Immaculata University, a master's degree in Educational Administration and Policy Studies from Temple University, and a bachelor's degree in Social Studies Education from Elizabethtown College. In addition to his professional positions and educational experiences, Dr. Hartman has earned the nationally recognized credential of Professional Human Capital Leader in Education (pHCLE) through the American Association of School Personnel Administrators (AASPA), serves as the co-chair of the professional development committee for AASPA, is a nationally certified principal mentor through the National Association of Elementary School Principals (NAESP), and is currently a policy fellow with the Pennsylvania Education Policy and Leadership Center (EPLC).

REVISITING MATHEMATICS LEARNING LOSS: EVIDENCE FROM LOWER ELEMENTARY STUDENTS AT THE SCHOOL-LEVEL

Joseph Anthes, Ed.D.
Craig Hochbein
Jihyun Kim, Ph.D.

PENNSYLVANIA EDUCATIONAL LEADERSHIP | VOLUME 41, NUMBER 1

ABSTRACT

More than a century of research exists on summer learning loss and summer learning programs. Although prior research demonstrates that summer learning loss takes a greater toll on math progress than reading (Cooper et al., 1996), research that examines the effects of summer learning programs on lower elementary math achievement is limited.

The current study analyzed a secondary database containing reading and math outcomes from two different summer learning programs for first and second grade students. The researcher capitalized on the elementary school's random assignment to conduct a series of experimental and quasi-experimental designs to compare fall benchmark assessments.

Outcomes from the study showed that neither summer program had a significant effect on math outcomes. Additionally, learning loss was more severe in math than in reading across all groups. Results from this study support further examination of K-2 summer math programs and greater urgency for addressing seasonal learning loss in mathematics.

INTRODUCTION

When students no longer have access to the consistent flow of school year instruction, resources, and structured activity, the rate of learning slows (Entwisle et al., 1997; Entwisle et al., 2001; Pitcock, 2018). However, as the rate of learning slows for all children during the summer months, the summer can also serve as a prime opportunity for children with academic deficits to catch up to their peers (Alexander, 2019; von Hippel, 2019). Additional learning opportunities may support school efforts to teach children to read proficiently by the end of third grade (Fiester, 2010) or in the case of math, prepare students with the pre-requisite skills needed access more advanced content and coursework in middle and high school associated with college and career readiness (Atanda, 1998). Failure to help students meet such crucial academic milestone can lead to higher rates of high school dropout, as well as negative social and economic consequences for students in adulthood (Atanda,1998; Fiester, 2010; Hernandez, 2011; NICHHD, 2000; Torgesen, 2002).

Summer reading loss is more severe for low-income students than their middle to upper-income peers. On the contrary, research shows that all students, regardless of socio-economic status, lose math skills during the summer (Cooper et al., 1996). In fact, Cooper et al. (1996) reported that all students lost a minimum of one month of math skills and averaged 2.6 months lost. Despite evidence that suggests an urgent need to combat summer math loss, research on lower elementary summer math programs is limited.

The purpose of this paper is to improve academic outcomes for students and close achievement gaps in math by focusing on summer math learning loss at the lower elementary level. The experimental research design assessed academic outcomes from a secondary database that included data from two summer learning programs attended by students from the same school population. To meet this purpose, the paper seeks to answer the following two research questions:

1. After controlling for spring benchmark scores, are there significant differences between groups in fall math benchmark scores?
2. How do summer changes in math benchmark scores compare changes in reading scores?

REVIEW OF LITERATURE

NAEP results show that only 38 percent of eighth graders demonstrate math proficiency (Ann E. Casey Foundation, 2019), an indicator that many students will be unable to access the higher-level math coursework in high school needed for success in college and post-secondary outcomes (ACT, 2008; Atanda, 1999; Cooney & Bottoms, 2003; Byun et al., 2015; Royster et al., 2015; Schneider et al., 1997). At the elementary level, Quinn et al. (2016) found that although math gaps between children in the highest and lowest socio-economic quintiles narrowed during kindergarten, gaps increased over the summer following kindergarten.

Like reading, mitigation of math learning difficulties requires high quality instruction in the early grades. Along with increased competence and fluency with basic addition and subtraction facts, young children either develop, or fail to develop, a basic number sense at an early age (Gersten & Chard, 1999). Number sense, or the ability to use numbers flexibly (Boaler, 2015), serves as the foundation for all higher-level math (Feikes & Schwingendorf, 2008). Developing number sense is a complex construct that when absent or impaired, may lead to longer-term mathematical difficulties (Berch, 2005). It is acquired through formal instruction and requires a principled understanding of place value, of how whole numbers can be composed and decomposed, and the meaning of basic operations (Greeno & Collins, 2008).

Gersten and Chard (1999) argue that number sense is analogous to phonemic awareness when developing foundational competencies necessary for future academic success. Whereas phonemic awareness relies heavily on auditory processing, number sense does not. However, the development of number sense may prevent the over identification of students with learning disabilities in math and assist in a more accurate identification and better instructional support of those that do have a math disability (Gersten & Chard, 1999).

On the other hand, math computation relies heavily on instruction and practice (Heyns, 1978; Wintre, 1986). Unlike reading, math may be less embedded in a family's regular household practices, leading math skills to decline more rapidly than reading skills during the summer (Pitcock, 2018), consistent with Cooper et al.'s 1996 meta-analysis that showed larger summer losses for math computation than for reading. Subsequently, learning loss in math contributes to an instructional and curricular redundancy more prevalent in United States classrooms than in international comparisons (Polikoff, 2012). Based on these findings, Cooper et al. (1996) recommended that summer programming target math skills.

However, recent evaluations of early elementary summer math programs are scarce at best. Summer reading program evaluations tend to be more prominent within the body of research, even though math learning loss is found to be more severe (Cooper et al., 1996). Just as students stand to benefit from additional instruction in foundational reading skills, summer instruction can also support foundational math skills and strengthen conceptual understandings. Supporting early elementary students' ability to conceptualize and use numbers flexibly contributes toward meeting proficiency on grade-level standards and strengthens their abilities as they advance toward more advanced mathematical thinking and tasks.

Unlike with reading, few summer math program evaluations exist that include students between kindergarten and second grade. The tendency for summer math evaluations to focus primarily on remedial programs in the intermediate and middle grades (Jacob & Lefgren, 2004; Mariano & Martorell, 2013; Matsudaira 2008) suggests that schools may overlook the summer as an opportunity to improve early elementary students' math skills. A possible explanation for this trend may be that state testing typically begins

in the middle of elementary school. The tendency for summer math programs to focus on students who fail to meet progress may be indicative of schools' and districts' response to meeting accountability standards on achievement tests.

The limited available evidence on the efficacy summer math programs is, however, encouraging. In 2014, McCombs et al. authored the largest randomized control trial on summer learning to date as part of five-volume series funded by the Wallace Foundation and conducted as part of the RAND Corporation. The purpose of the study was to determine whether voluntary, district-run summer programs improve academic achievement. The study consisted of 5,639 students from third to fifth grade across five United States metropolitan areas: Boston, Massachusetts; Dallas, Texas; Duval County, Florida; Pittsburgh, Pennsylvania; and Rochester, New York. In total, 3,194 students were assigned received treatment by participating in a summer learning intervention that included both reading and math.

This robust study ultimately determined that programs had positive effects on near term math achievement and that math outcomes were associated with better attendance and more instructional time. Despite the strength of this study, participating students had just completed the third grade, and thus the outcomes provided little evidence to support summer interventions for early elementary students. Due to dearth of evidence available, the need for experimental research on elementary summer math programs, especially those during students first three years of schooling, is critical toward eliminating current educational inequities (Quinn et al., 2016).

METHODS

This experimental study seeks to evaluate differences in summer reading and math progress within a small population of mostly low-income, urban elementary students. The researcher used a secondary database analysis from one elementary school that contained reading and math data for first and second grade students from the spring and fall of 2019. The analysis includes three experimental groups from the summer of 2019: students who declined to participate (DP), students who participated in the District Title I (DT1) summer program that primarily focused on reading, and students who participated in the Community School (CS) summer program that focused on both reading and math. All eligible students were randomly assigned to either the DT1 or CS program. Following random assignment students and families self-selected whether to participate in their assigned program. Those who declined to participate consented to be in the DP group.

The study utilized a randomized control trial to evaluate the effectiveness of an early elementary summer math intervention. Additionally, math outcomes were compared to the sample's summer reading outcomes from an additional quasi-experimental research design and analyzed using descriptive and inferential statistics. To investigate whether the school's math program met its intended purpose, as well as to compare student reading and math outcomes, the study sought to answer the following two research questions:

1. After controlling for spring benchmark scores, are there significant differences between groups in fall math benchmark scores?
2. How do summer changes in math benchmark scores compare changes in reading scores?

Population and School Context

Main Street Elementary School (pseudonym MSES) is a small, Title I K-5 elementary school. The school enrolls approximately 250 students annually and has two classes for each grade level. The student population is socioeconomically and racially diverse (Table 1).

Table 1

Spring 2019 Demographic Information

	N	F%	FRL%	EL%	SE%	NW%
SCHOOL POPULATION	252	46.6	77.1	5.1	14.6	77.2
ELIGIBLE POPULATION	82	42.7	69.5	7.3	2.4	56.1
POPULATION FINAL SAMPLE	54	50.0	70.3	0.5	3.7	55.5

Note: N = number of students, F = Female, FRL = Free/Reduced Lunch, ELL = English Learner, SE = Special Education, NW = Non-White. School demographic information is publicly available https://futurereadypa.org/.

MSES is an open-concept building and a community school, a model in which the school also serves as a community hub to local families and students. Community partners work with the school to integrate resources and services intended to provide basic needs and empower families. A local university serves as MSES's lead partner, and is located only one block from school. The vast majority of MSES students walk to school and live within one mile of the school. Despite the high ratio of families qualifying for free and reduced lunch, there are no subsidized housing units within the school's territory. The study's final sample includes slightly lower rates of students from historically marginalized populations.

Sample

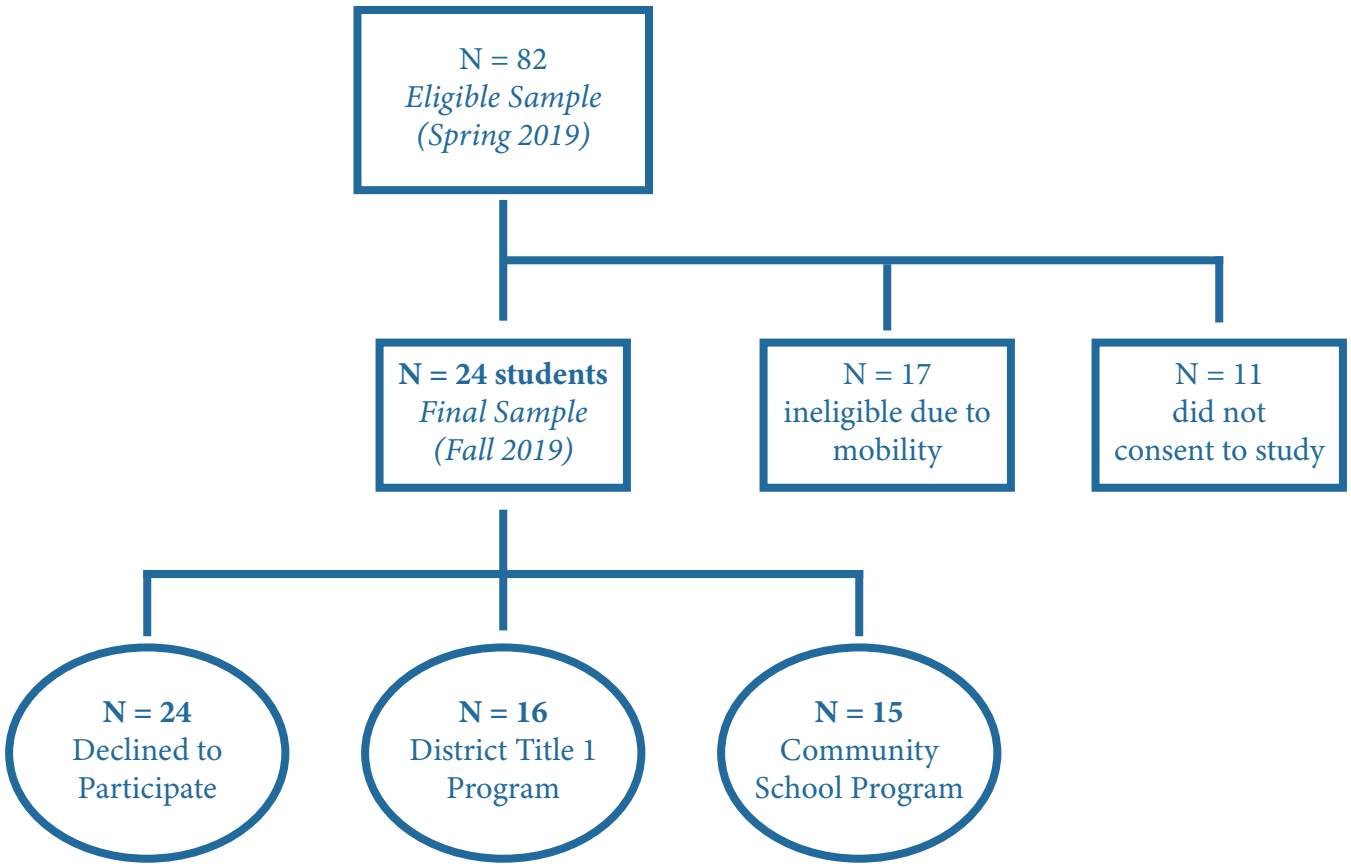
A non-probability, convenient sampling technique was employed within MSES. Approval was required from the university's Institutional Review Board before data collection could begin. Due to the small population, all 82 eligible first and second grade students were invited to participate in the study and one summer program, rather than randomly select students to participate. Students were randomly assigned at the population level at the onset of the study. The method of random assignment was chosen to allow families to feel informed about both the study and summer learning program participation. Random assignment began with each eligible student being assigned a participant number. Through the use of an online random assignment generator, students were assigned to either the DT1 or CS summer learning program.

The initial eligible population consisted of 82 students in spring 2019. The final sample consisted of 54 students in the fall and represented 83 percent of the eligible student population by fall 2019. The final sample (N = 54) included students who declined to participate in a summer learning program, but whose families consented to be in the study (N = 23), as well as students who participated in either the DT1 program (N = 16)

or CS program (N = 15). Seventeen students were ineligible by the fall due to summer mobility or changes in enrollment status. Eleven families chose not to participate in the study. Figure 1 details changes that occurred between initial sampling in the spring and the final sample that remained by the fall.

Figure 1

Final sample and experimental groups following random assignment, consent, and attrition related to mobility



The final sample's (N = 54) demographic profile was very similar to that of the eligible sample of first and second grade students (N = 82) The final sample had a slightly higher ratio of female students compared to the eligible sample, a slightly higher ratio of special education students, and a lower ratio of ELs (Table 1). However, the eligible sample and final sample differed by less than one percent in the ratios of student that identified as a race other than white and students qualifying for Free or Reduced Lunch.

DATA COLLECTION

Data was collected between May and September of 2019 from school databases. Data includes available student demographic information and both spring and fall reading and math benchmark results. Reading outcomes were assessed using Acadience Reading spring and fall benchmark assessments, and math outcomes were assessed using easyCBM math spring and fall benchmark assessments. In addition, summer program attendance was tracked and included in the dataset for all 31 summer participants.

EXPERIMENTAL CONDITIONS

DT1 Reading Instruction

Students from MSES were invited to attend the DT1 program that operated for five days a week for five consecutive weeks from June 17th to July 26th. The program had a total of 26 planned instructional days. Students attended programming for approximately five hours each day (Table 2).

Reading instruction consisted of lessons from the district's elementary core curriculum, Wonders, from McGraw Hill. Reading instruction took place for approximately 2 hours and 20 minutes each day. Eligible summer curriculum ranged from lessons in the final unit of Wonders that were not taught before the end of the regular school year to lessons in the first unit of the next grade's curriculum.

CS Reading Instruction

The CS program operated for three days a week for a duration of four weeks from July 15th to August 7th for a total of 12 instructional days. Student received reading instruction for approximately 2 hours and 20 minutes each day (Table 2).

Reading instruction focused on individual student intervention on targeted reading skills aligned to principles used in multi-tiered system of support (MTSS) intervention systems. In collaboration with university graduate students, the lead professor used assessments to guide small group instruction and individual instruction based on individual student needs.

CS Program Math Instruction

Math instruction focused primarily on number identification, number fluency, quantity discrimination, pattern and missing number identification, addition, subtraction, and word problem solving based on grade-level standards. Instruction was delivered through whole group and small group learning activity centers. Learning materials were developed and curated by teachers from online resources and the school's K-2 curriculum, *Investigations* by Pearson (Table 2). Teachers also utilized manipulatives such as snap cubes and flash cards. In addition, the principal at MSES worked with the university to connect school-provided devices to the university's wireless network so that teachers and students had access to district math applications and websites. Math teachers in the CS program were encouraged to use creativity, cooperative learning, and incorporate game-based learning to support student engagement in math and encourage continued summer program participation.

Table 2

MSES Summer Programs Comparison

DESCRIPTOR	DISTRICT TITLE I PROGRAM (DT1)	COMMUNITY SCHOOL PROGRAMS (CS)
DATES	JUNE 17 - JULY 25	JULY 15 - AUGUST 7
DURATION	6 WEEKS 26 TOTAL DAYS	4 WEEKS 12 TOTAL DAYS
DAILY TIME	5 HOURS	5.5 HOURS
TOTAL TIME	130 HOURS	66 HOURS
LOCATION	DISTRICT BUILDINGS	UNIVERSITY CAMPUS
STAFFING	MSES TEACHERS	GRADUATE STUDENTS MSES TEACHERS
PROGRAM ADMINISTRATION	GRANTS SUPERVISOR DISTRICT SITE COORDINATOR PRINCIPAL CS COORDINATOR	UNIVERSITY PROFESSOR PRINCIPAL INTERN PRINCIPAL CS COORDINATOR
READING INSTRUCTION	CORE CURRICULUM	MTSS INTERVENTION
MATH INSTRUCTION	NONE	TEACHER CURATED CORE CURRICULUM
ATTENDANCE	INCENTIVES, TRACKING	INCENTIVES, TRACKING

ANALYSIS

Research Question 1: After controlling for spring benchmark scores, are there significant differences between groups in fall math benchmark scores?

The first research question will compare fall math outcomes from each of the three groups to determine if differences exist between students who had no formal summer learning experience compared to students in each of the two treatment conditions. To evaluate general effects of summer program participation on students, participants from the DT1 program are included in this analysis to determine if participation in the district program, which primarily focused on reading, results in any differences in fall math outcomes compared to students who did not participate in any program. DP students did not receive reading or math instruction in either school-based program.

First, general descriptive statistics will be analyzed. In addition, I will compare the means of students meeting grade level cut points on each measure. Finally, a multiple regression model will be used to further analyze math outcomes while controlling for grade level differences and pre-test differences between groups using the following equation:

$$Y_i = \beta_0 + \beta_1 Pre\text{-}Measure_i + \beta_2 Group_i + \beta_3 Grade_i + e_i$$

where Yi is the fall benchmark outcome of student i. β_0 represents the intercept of the slope on the Y-axis. $\beta_1 Pre\text{-}Measure_i$ is the spring benchmark outcome that a student i participated in. $\beta_2 Group_i$ indicates the

experimental group that students participated in. This research question examined differences between students that participated in either summer program and students who declined to participate. The two categories for $\beta_2 Group_i$ will be transformed into two dummy variables for the analysis: dummy variable 1 (DT1 = 1; DP = 0; CS = 0) and dummy variable 2 (DT1 = 0; DP = 0; CS = 1). The DPs will serve as the group in which all other groups are compared to. $\beta_3 Grade_i$ is a control variable for the differences between student grades (first or second grade). ei is the random error term.

The model will analyze whether differences in math outcomes between DPs and the students in each of the summer programs while controlling for grade level differences and spring benchmark performance. Participation was optional and the DP group was self-selected. Therefore, it is important to control for pre-test measures in an effort to reduce selection bias when analyzing data for this research question.

Research Question 2: How do summer changes in math benchmark scores compare changes in reading scores?

This question will compare fall reading and math outcomes to analyze changes and compare student learning between both disciplines. Significant outcomes from inferential statistical analyses will be compared with specific attention to the treatment group (CS Group). In addition, mean outcomes and changes in percent of students meeting proficiency cut scores will be compared across groups.

RESULTS

To answer research question one, easyCBM math fall benchmark outcomes were analyzed. Mean scores and the percent of students proficient on each measure are again reported for the total sample and each experimental group. Additionally, fall math outcomes are compared to spring math outcomes for the total sample and for each experimental group. Following the descriptive overview, outcomes from the multiple regression model are assessed to determine the strength of the relationship between fall math outcomes and experimental group affiliation, while controlling for the effects of grade level and spring math outcomes.

Descriptive Statistics

The average easyCBM Math fall benchmark score for the entire sample was 58.74% correct (SD = 15.87). Scores ranged from 34% to 100% correct with a median score of 60% correct. Each group demonstrated losses between the spring and the fall (Table 3) with the average loss for all students being nearly 25% between the fall and spring benchmark assessments. While losses for all groups were similar and substantial, the DP group experienced the lowest mean learning loss (M = -23.3, SD = 16.11), followed by the CS program (M = -24.86, SD = 14.40) and finally the DT1 group (M = -27.44, SD = 14.79).

Table 3

easyCBM Math – Descriptive Statistics Overview

DESCRIPTIVE	DP	DTI	CS	TOTAL SAMPLE
TOTAL (N)	23	16	15	54
MEAN				
SPRING	87.17	78.88	83.53	83.70
FALL	63.87	51.44	58.67	58.74
MEAN CHANGE	-23.30	-27.44	-24.86	-24.96
SD				
SPRING	9.99	16.03	11.79	12.76
FALL	16.11	14.79	14.40	15.87
PROFICIENT (N)				
SPRING	22	10	12	44
FALL	19	8	10	37
% PROFICIENT				
SPRING	95.65	62.50	80.00	81.48
FALL	82.60	50.00	66.67	68.52
% CHANGE	-13.05	-12.50	-13.33	-12.96

All three groups had fewer students meet the proficiency cut score in the fall on the easyCBM math benchmark than did in the spring. 83% of students in the DP group were proficient on the easyCBM math benchmark assessment, whereas only 67% of CS program students were proficient and only 50% of DT1 students. However, all groups experienced similar losses in the percent of student that performed at benchmark between the spring and fall. The CS group experienced a greater loss amongst first grade students, as 25% fewer students (n = 2) failed to meet the math benchmark while there was no change in second grade.

Multiple Regression

A multiple regression model was used to perform a more robust analysis of the dataset and control for the effect that additional variables, including grade level and spring scores, had on fall outcomes. The multiple regression model was defined by the following equation:

$$Y_i = \beta o + \beta_1 Pre\text{-}Measure_i + \beta_2 Group_i + \beta_3 Grade_i + e_i$$

Spring easyCBM math scores and grade level were entered as independent variables. Two dummy variables were created for the experimental groups: dummy variable 1 (DT1 = 0; DP = 1; CS = 0) and dummy variable 2 (DT1 = 1; DP = 0; CS = 0). The first model tested whether the relationship between fall outcomes was different for students in either summer the DT1 or CS summer program compared to those who did not participate (DP). The results indicate that participation in the DT1 program was not a significant predictor of fall easyCBM scores b = -0.16, t(53) = -1.42, p = .16. Additionally, participation in the CS program was

not a significant predicted of fall easyCBM math scores b = -0.07, t(53) = -1.42, p = .55. Grade level was not a significant predictor of fall easyCBM scores b = -0.04, t(53) = -0.34, p =.73, but spring scores did significantly predict fall scores b = 0.66, t(53) = 6.25, p = .00 (Table 4). Based on results from the multiple regression analysis, neither the CS Program's summer math intervention nor the general effects of participation in the DT1 reading program resulted in a meaningful difference on student fall math benchmark scores.

Table 4

Multiple Regression Results Between Fall easyCBM Math Scores and Experimental Group Affiliation

SOURCE	B	SE B	β	*t*	*p*
CONSTANT	-5.34	12.34		-0.43	.67
GRADE	-1.10	3.22	-0.04	-0.34	.73
SPRING SCORE	0.81	0.13	0.66	6.25	.00
DP VS. DT1	-5.61	3.94	-0.16	-1.42	.16
DP VS. CS	-2.35	3.90	-0.07	-0.07	.55

NOTE: R=.71, R^2=.50, p<.05, *N*=54

The second model tested whether the relationship between fall outcomes was different for students in either summer the DP group or CS summer program when compared to those DT1 program as a reference group. The primary purpose for this model was to compare the outcomes between the DT1 program and CS program. Because participating summer students were randomly assigned, and because only students in the CS program received formal math instruction, this comparison represented a true randomized controlled trial.

Results showed that the second model explained a fair amount of variance in the outcomes (R2 = .50). The results showed that participation in the CS program (n = 15) was not a significant predictor of NWF CLS scores when compared to the DT1 program (n = 16), scores b = .90, $t(53)$ = .77, p = .45. Grade level was again not a significant predictor of easyCBM Math scores, but spring easyCBM Math scores continued to significantly predict fall scores b = .66, $t(53)$ = 6.25, p = .00 (Table 5). The multiple regression analyses revealed that math instruction in the CS program did not result in a significant effect on participating students' fall easyCBM math outcomes.

Table 5

Multiple Regression Results Between Fall easyCBM Math Scores and Experimental Group Affiliation

SOURCE	B	SE B	β	*t*	*p*
CONSTANT	-10.95	11.58		-0.95	.35
GRADE	-1.10	3.22	-0.04	-0.34	.73
SPRING SCORE	0.81	0.13	0.66	6.25	.00
DP VS. DT1	5.61	3.94	0.16	1.42	.16
DP VS. CS	3.26	4.26	0.09	0.77	.45

NOTE: R=.71, R^2=.50, p<.05, *N*=54

The second research question compared how spring to fall math changes compared to changes in reading during the period between spring and fall benchmark assessments. ANOVA and multiple regression were used to determine whether summer interventions resulted in statistically significant results. A descriptive analysis examined both changes in mean scores and the percent of students that remain above the proficiency cut score for each measure at both spring and fall intervals.

The CS program had a statistically significant effect on NWF CLS outcomes ($p < .05$). Despite the statistically significant findings for the CS group, the percent of students meeting benchmark cut points in the NWF CLS measure declined for the across groups (-14.81%), including the CS group, between spring and fall (-13.33%). Gains among already high performing students may have attributed to the increased group mean statistically significant results in the CS group. Neither the DT1 and CS summer programs had a statistically significant effect on any of the other fall reading measures (Table 6).

Table 6

Analysis and Significance for Each Reading and Math Measure Included in the Study

MEASURE	STATISTICAL ANALYSIS	SAMPLE (N)	STATISTICAL SIGNIFICANCE ($p < .05$)	SIGNIFICANT EXPERIMENTAL INTERVENTION
NONSENSE WORD FLUENCY CORRECT LETTER SOUNDS	MULTIPLE REGRESSION	54	SIGNIFICANT	CS GROUP
NONSENSE WORD FLUENCY WHOLE WORDS READ	ANCOVA	30	NOT SIGNIFICANT	NONE
PHONEME SEGMENTATION FLUENCY	ANCOVA	24	NOT SIGNIFICANT	NONE
ORAL READING FLUENCY – WORDS READ	ANCOVA	30	NOT SIGNIFICANT	NONE
ORAL READING FLUENCY – ACCURACY	ANCOVA	30	NOT SIGNIFICANT	NONE
EASYCBM MATH	MULTIPLE REGRESSION	54	NOT SIGNIFICANT	NONE

In comparison, a multiple regression analysis showed that the CS program, which was the only summer program that included a math intervention, had a non-significant effect on fall math outcomes (Table 6). Mean differences between fall and spring easyCBM math benchmark scores demonstrates that all groups declined in math skills despite any intervention (Figure 2). Scores on the fall benchmark assessment declined in the CS group (M = - 27.44) as well as across the total sample (M = -24.96), as did the percent of students meeting benchmark cut points compared to the spring (CS = - 12.5%; Total Sample = 13.33%). With the exception of NWF CLS, the decline in math scores during the summer is greater than any of the other Acadience Reading measures administered. However, the percent of students that remain above the proficiency cut point for fall easyCBM math remained higher than the ratio of students above the oral reading fluency cut points for both fluency and accuracy (Figure 3).

Figure 2

Bar chart of spring and fall easyCBM math performance across each experimental group and total sample

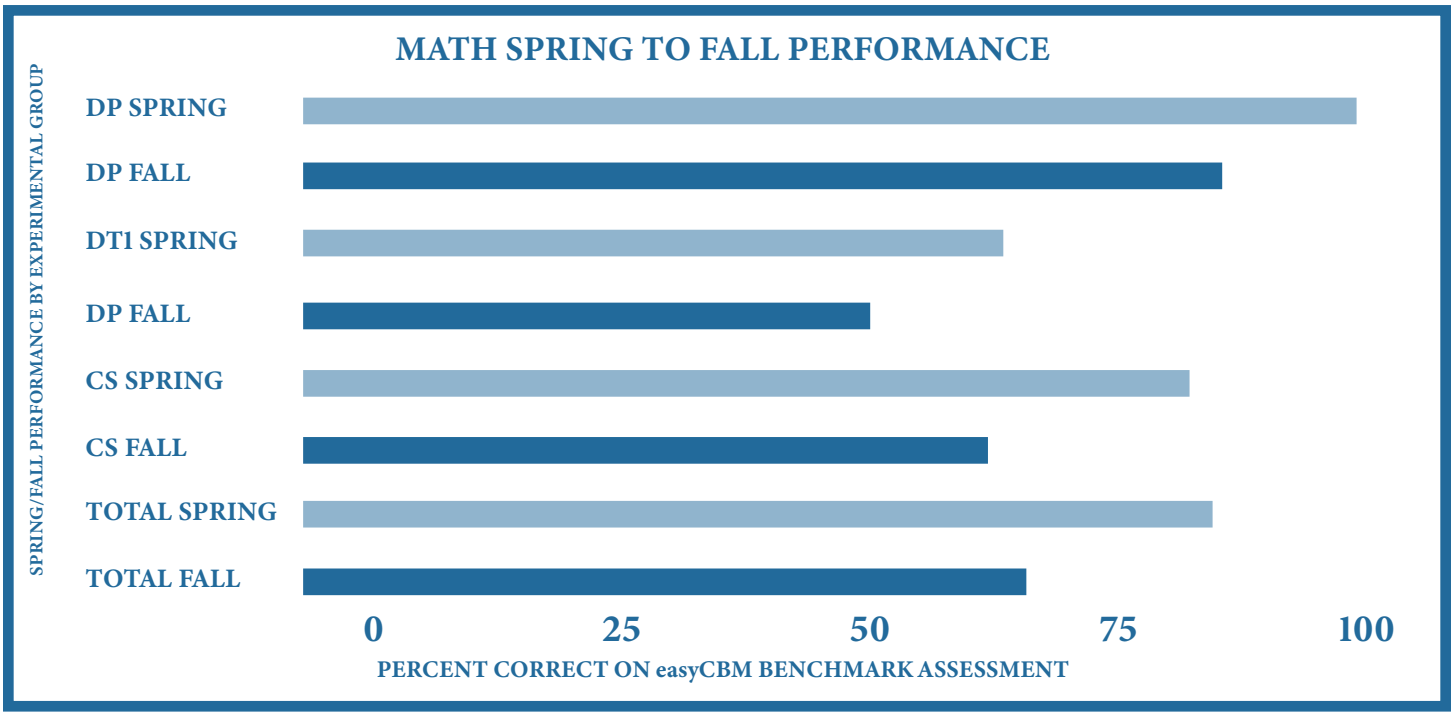
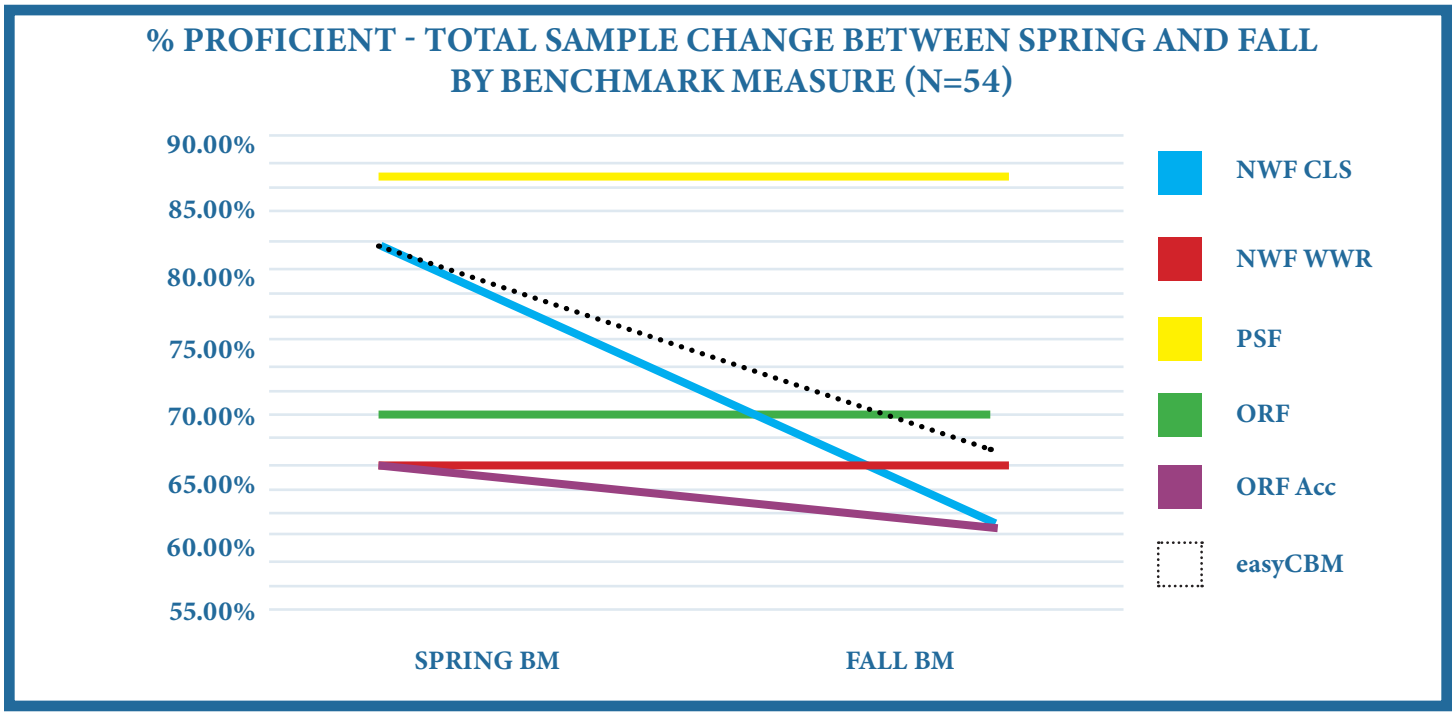


Figure 3

Line graph of the total percent change in proficiency between spring and fall for each reading and math measures for the study sample



DISCUSSION

MSES was in a unique position to provide two different summer learning opportunities to first and second grade student allowing for a randomized control trial and subsequent analysis of student reading and math outcomes for those that received. Results from this study support the original claim by Harrison Cooper et al. (1996) that students suffer greater math losses than reading. Whereas students in the CS program saw a range of maintenance, small losses, and even modest gains in reading skills, math learning loss showed to be more severe and universal for all students.

Math loss appear more pronounced than overall reading losses. With the exception of NWF CLS, the percent of students remaining at or above Acadience Reading benchmark cut points appeared to remain consistent. The decline in ORF accuracy scores was less dramatic that of easyCBM math scores and NWF CLS scores, while fluency scores held steady, despite their already low start point in the spring. However, the use of multiple reading measures compared to a single math measure requires further investigation to provide a more conclusive assessment.

As a result of overall concern for math learning loss, future programmatic efforts should consider increased time and multiple assessment measures for mathematical progress and effectiveness of math interventions. The CS program offered 12 days of instruction over 4 weeks for a total of only 2 hours of math instruction daily. Considering the known magnitude of math learning loss, summer programs should seek to provide increased learning opportunities for students. Additionally, the type of assessment, as well as the frequency and the length between program completion and benchmark administration, must be considered. easyCBM math is a comprehensive measure that provides a benchmark of student learning across all common core math domains, whereas Acadience Reading measures serve as indicators of specific grade level reading skills. Therefore, a limitation of this study is that math is discussed in less specific detail than reading. Future studies should consider specific, targeted math instruction with specific assessment to assess the targeted skill. In addition, progress monitoring measures across programs can assist in determining whether summer interventions have a positive impact during programming, in addition to how learning during summer programs translates at the start of the school year.

Future research on summer learning programs at any scale should also continue to examine K-2 math programs and outcomes. Research that examines the effects of summer and summer learning programs on K-2 math achievement is limited. Outcomes from the current study showed that learning loss was more severe in math than in reading across all groups. Additionally, the short duration of the math program, coupled with a gap between program completion and the fall benchmark, showed that the CS math program did not have an effect on student outcomes. Rather than wait until students require remediation in upper elementary, middle, or high school, schools should seek every opportunity to help students achieve grade level proficiency in the lower grades. Future research that further explores the specific patterns for math learning, as well as methods for successful summer math intervention, can help educators develop best practices for how to best use the summer months to support students and increase opportunities for learning in math.

CONCLUSION

The current study is significant for several reasons. First, random assignment within field settings is more difficult than in laboratory settings and therefore, far less common (Cook et al., 1979) and rare when evaluating summer learning programs (Kim & Quinn, 2013). Additionally, the unique research design accounted for the role of self-selection by including DPs as an additional comparison group. A third strength is that despite non-significant statistical outcomes, this study is one of few that evaluated and reported the effects of summer math instruction on students younger than third grade.

Finally, a unique strength of the study is that it is one of few evaluations that assessed the impact of summer learning programs on a school-level population. The dearth of evaluations that investigate to what extent summer programs contribute to school improvement may be for good reason. Studying a school population comes with challenges of smaller samples, and therefore more negative effects of attrition or student mobility. Summer programs are challenged by enrolling students and recruiting teachers during summer vacation, as well as scarce resources (McCombs et al., 2011; Von Drehle, 2010). It is likely that these same challenges are exacerbated at the school-level prevent schools and researchers from attempting to study the effects of summer interventions on school populations. The inclusion of studies on school-level populations is important, as it is individual schools that bear the burden of accountability and improved student performance. Therefore, evaluating whether summer learning programs can be a viable tool to improve achievement is a critical research endeavor that until this study, has been underexplored.

Math may be more difficult to address than reading. Evidence of disrupted math progress from K-2 summer learning evaluations suggest a reconsideration, or at the least, an awareness of the detrimental effects that the summer has on elementary math progress. School and district leaders should reallocate time and resources, as well as consider instructional and teacher quality as it pertains to math, when designing programs that aim to prevent learning loss and help promote continued student growth in math.

REFERENCES

ACT. (2008). *The forgotten middle: Ensuring that all students are on target for college and career readiness before high school.* Iowa City, IA. https://www.act.org/content/dam/act/unsecured/documents/ForgottenMiddle.pdf

Alexander, K. L. (2019). Summer learning loss sure is real?. *Education Next, 19*(4).

Atanda, R. (1999). Gatekeeper courses. *National Center for Education Statistics, 1*(1), 33.

Annie E. Casey Foundation. (2019). *KIDS COUNT Data Center: Pennsylvania Indicators.* https://datacenter.kidscount.org/data#PA/2/0/char/0

Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38*(4), 333-339.

Boaler, J. (2015). *Mathematical mindsets: Unleashing students' potential through creative mathematics, inspiring messages and innovative teaching.* John Wiley & Sons.

Byun, S. Y., Irvin, M. J., & Bell, B. A. (2015). Advanced math course taking: Effects on math achievement and college enrollment. *The Journal of Experimental Education, 83*(4), 439-468.

Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Houghton Mifflin.

Cooney, S., & Bottoms, G. (2003). Middle Grades to High School: Mending a Weak Link. Research Brief. https://www.semanticscholar.org/paper/Middle-Grades-to-High-School%3A-Mending-a-Weak-Link.-Cooney-Bottoms/12412518337312fd203baaa498c8f8fb7fe21bf5

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 66(3), 227-268.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (1997). *Children, schools, & inequality*. Westview Press.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2001). Keep the faucet flowing summer learning and home environment. *American Educator,* 25(3), 10-15.

Feikes, D., & Schwingendorf, K. (2008). The importance of compression in children's learning of mathematics and teacher's learning to teach mathematics. *Mediterranean Journal for Research in Mathematics Education*, 7(2).

Fiester, L. (2010). *Early warning! Why reading by the end of third grade matters. Kids count special report. Annie E. Casey Foundation.*

Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education, 33*(1), 18–28. https://doi.org/10.1177/002246699903300102

Greeno, J. G., & Collins, A. (2008). Commentary on the final report of the National Mathematics Advisory Panel. *Educational Researcher, 37*(9), 618-623.

Hernandez, D. J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation. Annie E. Casey Foundation.*

Heyns, B. (1978). *Summer learning and the effects of schooling.* Academic Press.

Jacob, B.A., & Lefgren, L. (2004). Remedial education and student achievement: A regression- discontinuity analysis. *Review of Economics and Statistics, 86*(1), 226–244.

Kim, J. S., & Quinn, D. M. (2013). The effects of summer reading on low-income children's literacy achievement from Kindergarten to Grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research, 83*(3), 386-431.

Mariano, L. T., & Martorell, P. (2013). The academic effects of summer instruction and retention in New York City. *Educational Evaluation and Policy Analysis, 35*(1), 96-117.

Matsudaira, J. D. (2008). Mandatory summer school and student achievement. *Journal of Econometrics, 142*(2), 829-850.

McCombs, J. S., Augustine, C. H., Schwartz, H. L., Bodilly, S. J., McInnis, B., Lichter, D. S., & Cross, A. B. (2011). *Making summer count: How summer programs can boost students' learning.* RAND Corporation.

McCombs, J. S., Pane, J. F., Augustine, C. H., Schwartz, H. L., Martorell, P., & Zakaras, L. (2014). *Ready for Fall? Near-Term Effects of Voluntary Summer Learning Programs on Low-Income Students' Learning Opportunities and Outcomes.* RAND Corporation.

National Institute of Child Health and Human Development (NICHHD). (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* (NIH Pub. No. 00 – 4754). Washington, DC: U.S. Government Printing Office.

Pitcock, S. (2018). The case for summer learning: why supporting students and families all year is vitally important. *American Educator, 42*(1), 4.

Polikoff, M. S. (2012). The redundancy of mathematics instruction in US elementary and middle schools. *The Elementary School Journal, 113*(2), 230-251.

Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. *Educational Researcher, 45*(8), 443–453.

Quinn, D., & Polikoff, M. (2017). Summer learning loss: What is it, and what can we do about it. *Brookings Institution*. https://www.brookings.edu/research/summer-learning-loss-what-is-it-and-what-can-we-do-about-it/

Royster, P., Gross, J., & Hochbein, C. (2015). Timing is everything: Getting students back on track to college readiness in high school. *The High School Journal, 98*(3), 208-225.

Schneider, B., Swanson, C. B., & Riegle-Crumb, C. (1997). Opportunities for learning: Course sequences and positional advantages. *Social Psychology of Education, 2*(1), 25-53.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology, 40*(1), 7-26.

Von Drehle, D. (2010). The case against summer vacation. *Time Magazine, 176*(5), 36-42.

Von Hippel, P. T. (2019). Is summer learning loss real?. *Education Next, 19*(4), 9-14.

Wintre, M. G. (1986). Challenging the assumption of generalized academic losses over summer. *The Journal of Educational Research, 79*(5), 308-312.

ABOUT THE AUTHORS

Joseph Anthes, Ed.D., is a K-12 Curriculum Supervisor in the Bethlehem Area School District. Anthes also serves an adjunct professor of education, working with pre-service teachers in the Lehigh Valley. He earned a BA in history from Gettysburg College, an MAT from the College of New Jersey, a MS in Educational Administration from the University of Scranton, and a doctorate in Educational Leadership from Lehigh University.
anthesj1@gmail.com; janthes@basdschools.org

Craig Hochbein is an Associate Professor in the Educational Leadership Program at the Lehigh University College of Education. His research examines school leader time use, declining academic achievement within school, and the effectiveness of policies intended to improve school performance. Hochbein also serves as an assistant editor for the journal *School Effectiveness and School Improvement*.

Jihyun Kim, Ph.D., is an assistant professor in the Department of Education at Sungshin Women's University in Seoul, Korea. Jihyun Kim received her Ph.D. degree in educational policy from Michigan State University. Her research interest includes teaching quality, teacher evaluation, policy evaluation, policy implementation, and principals' leadership. She has presented her work nationally and internationally and published in *American Educational Research Journal, Teachers College Record, American Journal of Education, and Teaching and Teacher Education.*

PREPARING PRE-SERVICE SPECIAL EDUCATORS TO CO-TEACH FINANCIAL LITERACY

Melinda S. Burchard, Ph.D.
Alyssa Amels
Karissa Hagar
Carol Buckley, Ed.D.

PENNSYLVANIA EDUCATIONAL LEADERSHIP | VOLUME 41, NUMBER 1

ABSTRACT

Special educators need preparation to co-teach personal finance content. This study investigated outcomes of a training module teaching pre-service special educators to co-teach personal finance. Researchers examined interactions between self-efficacy for interventions with outcomes of gains in knowledge, budgeting skills, and quality of lesson design for teaching personal finance content. Results demonstrated positive correlations between the quality of lessons designed by the pre-service teachers and self-efficacy. Student gains in personal finance knowledge reached a very large effect size of Cohen's d= 1.14. Qualitative analysis revealed three common themes of "wants versus needs", "decisions impact finances", and "planning for the unpredictable".

Keywords: financial literacy, personal finance, special education, teacher education

INTRODUCTION

Financial Literacy

Educators are prioritizing secondary instruction in financial literacy. Currently, twenty-one U.S. states require high school content addressing financial literacy, with five states requiring stand-alone courses in personal finance (Rosenbaum, 2020) and other states encouraging such courses (Public School Code, 2019). Analyzing national data about student financial aid, researchers found that students who took mandated financial literacy courses made college borrowing decisions that appeared wiser for long-term debt. They concluded that such required courses appeared to support wiser financial decision-making (Stoddard et al., 2018). In the National Financial Capability Study (2018), most Americans failed to answer more than 50% of a short list of questions about personal financial literacy.

Financial Literacy Education in Pennsylvania

In the Commonwealth of Pennsylvania, Act 91 of 2019 adapted high school graduation requirements to include the option to apply one credit from a course in personal finance. Those requirements allow for such a personal finance course to be taught by teachers of mathematics, business education, social studies (specifically economics), or family and consumer sciences (Public School Code, 2019).

Financial Literacy for Students with Disabilities

Focusing on needs of students with high incidence disabilities, Henning & Johnston-Rodriguez (2018) evaluated personal finance curricula for accessibility and universal design for learning. They found five curricula with promise for use with students with high incidence disabilities, though none were strong in all components evaluated such as "multiple forms of expression" or "multiple forms of engagement". Further, they stressed the importance of addressing disability-related economic injustice and inequities (Henning & Johnston-Rodriguez, 2018). Barczack (2019) stresses that when teaching independent living skills of personal finances to students with intellectual disabilities, educators must sequence from explicit instruction in special education to authentic practice in community settings.

Supporting Transition to Independence with Financial Literacy Training

The Individuals with Disabilities Education Act, IDEA (Individuals with Disabilities Education Act, 2004), established requirements for how eligible students with disabilities are served. One specific requirement of that act is planning toward the successful transition for students into postsecondary training or education, postsecondary employment, and independent living. While IDEA requires such planning to begin at age 16 (Individuals with Disabilities Education Act, 2004), specific states may start such planning earlier. Support for independent living can include a broad range of skills such as hygiene, laundry, cooking, leisure skills, navigating transportation, and certainly managing money.

As school districts offer high school courses in personal finance, special educators should anticipate students with disabilities to enroll in such courses. Regardless of potential co-teaching delivery of high school personal finance courses, secondary special educators should logically expect to teach skills in financial literacy

as one component of supporting transition of students toward post-secondary independent living. Therefore, individuals pursuing teacher certification in special education would benefit from preparation to teach the math and reasoning of financial literacy content.

Special Education Teacher Preparation

Preparation for teaching in special education requires preparing students to meet numerous competencies, both in breadth and depth. Those requirements vary by state but most states align carefully with the Initial Preparation Standards by the Council for Exceptional Children, CEC, the leading professional organization in special education (CEC, 2015). For success in initial practice, special educators must be ready to teach across both academic and functional content and to apply specially designed techniques to meet unique needs of learners.

Educators must hold students in special education to high expectations. Co-teaching is one approach to delivering an appropriate education to children with qualifying disabilities. Successful co-teaching requires knowledge and skills in various models of such collaborative teaching, but also requires knowledge and skills in each content area in which a special educator co-teaches with a content expert. Importantly, a co-teacher needs to be more than a passive observer, an extra body in the classroom, but instead needs to add the expertise of designing individualized instruction for students with disabilities in the context of the co-taught classroom and the curriculum being taught (Friend, 2016; Rodgers & Weiss, 2019; Weiss & Glaser, 2019).

From 2012 to 2021, Pennsylvania Department of Education (PDE) certified special education teachers in either of two grade bands, grades PreK-8 or grades 7-12. In 2019, PDE revised special education guidelines to require teacher certification in across all grades PreK through 12th, effective January of 2021 (Pennsylvania Department of Education, 2019). Such extension of grade band for special education certification required curricular revisions to prepare pre-service special educators to teach and co-teach content across the full grade span, from early interventions to high school content.

Competencies Developed in Personal Finance Training

Preparing pre-service special educators to co-teach in personal finance is consistent with three CEC Initial Special Education Preparation Standards: 1) "Curricular Content Knowledge", 2) "Instructional Planning and Strategies", and 3) "Collaboration" (CEC, 2015). "Curricular Content Knowledge" includes developing knowledge in all content they might teach, including how to employ specialized techniques to support needs of individual students and adaptations to make content accessible. Co-teaching is a special education delivery model through which "...educators focus on integrating into daily lessons the special education strategies and techniques that will enable students to achieve the goals of their IEP [...individualized education plan]" (Friend, 2016, p 18). The standards for "Instructional Planning and Strategies" require competence in finding and evaluating evidence-based practices, supporting transition, critical thinking and

problem-solving. "Collaboration" includes the expectations to effectively collaborate, even co-teach, with others across varied settings (CEC, 2015). Through the module on personal finance, the faculty member required students to develop content knowledge appropriate for one high school course they may not have taken in their own prior learning, and to consider how to teach a narrow slice of that content using specific evidence-based teaching practices. The original plans for this module involved a requirement to co-teach their lesson with a high school teacher of personal finance. To adapt for pandemic-related protocols, the faculty adapted requirements. Students designed a video lesson that could be used by that high school teacher as a station or virtual engagement activity.

Self-Efficacy for Teaching

The specialized curriculum of special education courses integrated with strong foundational core curriculum in instructional design and core content develop self-efficacy for pre-service teachers to perform across content and specialized roles. Self-efficacy is defined as "...beliefs in one's capabilities to organize and execute the course of action required to produce given attainments" (Bandura, 1977). Teacher education research demonstrates correlations between self-efficacy for teaching, responsive flexibility, and resilience in teaching (Erdem & Demirel, 2007; Hoy et al., 2009).

Collaboration in Instructional Design

To address state certification changes, instructors at one university collaborated to revise one special education course preparing special educators to co-teach core content through high school grades. The university faculty specifically designed content to prepare pre-service teachers for the possibility of co-teaching personal finance.

In preparation to add this new module, three public school educators provided consultation and support to the university faculty. From one district, a math coach and a math teacher of personal finance advised on priority content knowledge for co-teachers and options for pre-service teachers to demonstrate co-teaching skills. From a second district, a teacher of business education shared access to lessons and assessments to support development of the lessons and assessments for the pre-service teachers. Each teacher also suggested texts to study. In addition, a second university faculty member provided guidance on development of the math skills to complement learning in a concurrent math methods course, and collaborated on assessment design.

The math intervention content of that course leading up to the personal finance module emphasized evidence-based math intervention practices, and co-teaching roles within special education. Instruction specific to personal finance covered broad topics such as budgeting and types of insurance, algebraic reasoning to solve financial challenges, and awareness of disability-related economic inequities.

PURPOSE OF THIS STUDY

Consistent with Boyer's model for the scholarship of teaching and learning (Boyer & Carnegie Foundation, 1990), the researchers investigated how the curriculum revisions worked. Specifically, the researchers investigated the following questions:

1) What, if any, interactions exist between self-efficacy, knowledge of personal finance, and quality of lesson design to co-teach personal finance?
2) How effective is the instruction on outcomes of knowledge and skills in personal finance?
3) What patterns exist in student reflections?

METHODS

The researchers employed a mixed methods design to explore their research questions. Prior to the study, they pursued and received approval from their Institutional Review Board, IRB.

Inclusion and Exclusion Criteria

Inclusion in this study required three criteria. First, all participants were university students, aged 18 or older. Second, inclusion required enrollment in a special education course about learners with high incidence disabilities, focused specifically upon students accessing the regular education curriculum. Finally, the researchers only included participants registered in one of three academic programs, a major pursuing teacher certification in special education, a major pursuing dual teacher certification in both special education and elementary education, or a minor in special education. The researchers excluded data if students were aged 17 or younger, not enrolled in the course, or not registered in one of the special education majors or minor.

The IRB approval allowed for testing for similarity and possibly combining data from two separate student cohorts who progressed through the unit in two different semesters. Pandemic-related social distancing protocols resulted in significant differences in the manner of instructional delivery. Therefore, the researchers determined that the two cohorts of students did not receive the same training. Thus, researchers evaluated only data from the second year of this study.

Assessment and Training Sequence

The faculty delivered this training in the math intervention module of an undergraduate course about teaching students with high incidence disabilities. Participants enrolled in the course in fall of their junior year, having all completed a pre-requisite course covering introduction to special education.

Pre-Assessments

At the start of the semester, students rated their self-efficacy in multi-tiered intervention practices using the Multi-Tiered Instruction Self-Efficacy Scale, MTISES (Barnes & Burchard, 2011, 2016). On the first day of the math intervention module, students also completed a pre-assessment of personal finance knowledge. The faculty graded this pre-assessment for participation only, but provided corrective feedback individually and through lessons.

Training

The faculty of this course delivered training over a series of weeks. Due to social distancing within a pandemic, students participated as members of two teams. One team attended class face-to-face while the other team attended class virtually, with equal days of face-to-face or virtual attendance on a rotating basis. All students accessed the same lessons, and engagement activities with pedagogy combined to meet needs of both virtual and in-person learners.

Because personal finance content is so broad, the faculty focused upon an introduction to selected big ideas of personal finance. Instruction covered information and corrective feedback for every topic assessed in the pre-assessment. Students learned about and discussed how most models of personal finance work well for middle class and up, but not as well for individuals living in poverty. Students extended discussions from an earlier module about social justice into this discussion. Students learned about various models or options for budgeting, debt management, savings, insurance, growing credit records, maintaining tax records, and protecting financial security. Simple skills development focused upon computing budgets and net worth, debt payment strategies, and compounding interest. The faculty also taught students to translate scenarios into word problems, then using word problem-solving strategies to solve the math. For each of those problems, the faculty required students to name the math problem types, similar to math intervention strategies covered earlier in the module. For example, in one scenario, a couple is computing anticipated monthly retirement income with one reporting annual income and the other reporting monthly incomes. Students must first "share" the one spouse's annual income across twelve months, before "combining" monthly incomes.

Active Engagement

The faculty engaged students through in-class problem-solving challenges embedded in each lesson, with every student called upon. Students addressed some challenges working with partners in small groups. Some active engagement simply required students to categorize expenses into budget categories. Other lessons required students to problem-solve using a specific strategies, or to compute the math from a personal finance scenario. Some emphasized naming the math required by a financial scenario.

Teachers' Budget Game

Students also played a budget game. Prior to the game, students answered survey questions choosing one job from teaching choices in urban, suburban, or rural, a job as an occupational therapist, or a job in museum education. Based upon those choices, students then chose living arrangements (buying, renting, and whether or not to have roommates), transportation options (buying new, buying used, or public transportation), and various common benefits beyond the package offered by their job. The instructor sent an "HR Letter" to each student recording their choices, indicating their gross and net pay, co-pays for various types of insurance, and reminding them of their other choices to inform their game play. Students all started the game with the same amount in a simulated savings account then played the budget game through 40 events. For each event, students rolled dice or flipped a coin with two possible outcomes. For example, one involved

an unexpected health event. Students who rolled an even number or flipped heads experienced a common cold that required purchase of over-the-counter cold remedy. Students who rolled an odd number or flipped tails required a doctor's visit with co-pay and prescription varying on the benefits choices made earlier. With each event, students recorded paychecks, school loan payments, rent, and varied other expenses and incomes into a spreadsheet. Students then submitted their spreadsheet and wrote reflections on the experience.

Post-Assessments

The training unit culminated in two assessments. Students again completed an assessment of personal finance knowledge, with every item the same as those asked in the pre-assessment quiz. The faculty graded this post-assessment of knowledge for accuracy. Each student signed up for one narrow personal finance topic, then designing an instructional video on that topic appropriate to teach a high school student with a learning disability. Assignment details for the instructional video included explicit teaching, modeling, and one way to actively engage the learners. The faculty graded quality of instructional video using a rubric.

Instrumentation

Researchers employed varied quantitative instruments to assess self-efficacy, knowledge, and skills. In addition, researchers qualitatively analyzed written reflections.

MTISES, Multi-tiered Instruction Self-Efficacy Scale

To assess self-efficacy, students completed self-ratings at both the beginning and end of the semester using the Multi-Tiered Instruction Self-Efficacy Scale, MTISES. Previous studies demonstrated that the MTISES works as a measure of self-efficacy and works with strong validity and reliability to measure needs for professional development and gains in response to training (Barnes & Burchard, 2010, 2011, 2016). In one study, the MTISES worked to measure interactions between self-efficacy and quality of IEPs, a legal document used to plan individualized services for children in special education (Burchard & Vargas, 2020). The MTISES also worked to assess interactions between self-efficacy and skills to find, evaluate and write about evidence-based teaching practices (Burchard & Myers, 2019) and how teachers attribute causes of learning outcomes with implementation of evidence-based teaching practices (Burchard et al., 2017).

The MTISES includes 28 items, each asking students to rate their need for professional development on a Likert Scale ranging from "I'll take anything" to readiness to help others. Each item asks students to rate need for professional development. For example, "How much professional development do you need to find research-based articles and/or books on practices relevant to specific educational needs of students?" or "How much professional development do you need to use data from appropriate assessment tools to clarify the specific problem for a struggling student?" (Barnes and Burchard, 2011, 2016).

Of the six MTISES subscales, five fit this specific study. Therefore, the researcher computed interactions with scores for overall self-efficacy for multi-tiered practices and for the subscale categories: evidence-based solutions, data-driven decision-making, engaging learners, implementing interventions, and collaboration.

Assessment of Personal Finance Knowledge

The course faculty collaborated with math education faculty to design a test of knowledge and skills. Students completed that test as a formative assessment at the beginning of the instructional module and again as a summative assessment at the conclusion of the module. The test included 18 multiple-choice questions, assessing knowledge of budgets, saving, records and credit scores. Half of the test items required declarative level reasoning to choose one correct answer. Five test items required application to a brief scenario. Four items required problem-solving.

Questions varied in complexity. One simple question asked students to identify whether to use gross pay or net pay in establishing their budget. One less complex question asked, "Which of the following is NOT a NEED according to budget experts?" with response options including items fitting wants, needs and savings options. Two items used the same list of financial details, one requiring students to compute assets, the second requiring students to compute liabilities. They computed split costs for a shared trip, compounding interest for a 401K account, monthly retirement income. The most complex problem required students to compare benefits of two job offers involving payment of school loans and grants and graduate school benefits. One item asked students, "What is this personal finance scenario asking me to DO? Meg and Alex are planning a monthly budget for their retirement. After taxes, Meg earns $1,000 per month from retirement. She earns an extra $726 each month from an investment. Alex earns $12,600 per year from an annuity. As a pair, what will they earn each month?" Instead of asking students to compute the answer, the answer options for that item required students to choose between the options of: divide then add, distribute, apply proportional reasoning, or add fractions.

Budget Spreadsheet and Reflections

Students played a budget game that involved two assessments, a spreadsheet and written reflections. For some events, students entered their net pay, paid their rent, or transportation payments. For other events, students either flipped a coin (heads or tails) or rolled dice (even or odd). That chance outcome directed their specific response to the event. For example, one item addressed moving expenses. If students rolled even or flipped heads, "Friends help you move. Buy pizza as a thank you. $30.00." If, however, students rolled odds or flipped tails, "Rent a truck to move your stuff. $120.00." For a health-related event, students who rolled even or flipped heads experienced "a common cold" with an over-the-counter remedy costing $6.78. On the other hand, students who rolled odd or flipped tails experienced a doctor visit with prescription with expenses varied depending upon the health insurance selections students made prior to the start of the game. In one income event, students received money in the mail from a relative with the flip of the coin or roll of the di determining how much. One event required students to make their own choice of whether or not to attend the wedding of a friend, just sending a gift for $50 or paying costly flight, hotel, and expenses to attend totaling $1,257.35.

The faculty graded the spreadsheet for accuracy in computations for each of the 40 events. Students submitted reflections through a discussion in the learning management system, which the faculty graded for

participation. The researchers analyzed those reflections qualitatively for common themes.

Rubric for Demonstration Videos

In grading students' video lessons on narrow topics of personal finance, the faculty rated performance using a rubric. Scoring emphasized alignment with standards, video quality, and demonstration of evidence-based practices (especially modeling as part of explicit teaching and active engagement with interactive materials). Students also earned points for quality of feedback to peers and participation on video demonstration day.

Qualitative Coding

Researchers analyzed one assessment qualitatively, the students' reflections on the budget game experience. The faculty recorded discussion text into a spreadsheet, removing names for confidentiality. Researchers then coded those texts to identify common themes in student reflections.

Data Analysis Software

Researchers compiled all quantitative data into Statistical Package for the Social Sciences, SPSS. They compiled qualitative data into EXCEL for coding of common themes.

RESULTS

Participants

Study participants were all students at a private university with just over 2,600 undergraduate students, and located in south central Pennsylvania (Messiah at a Glance, 2020). After applying inclusion and exclusion criteria, researchers analyzed data from twenty total participants who took the course in fall 2020. All twenty were traditionally aged university students. Most students took the course as a requirement for teacher certification in special education, one seeking certification in only special education, seventeen seeking dual certification with both special education and elementary education. Two students took the course toward a minor in special education, one pursuing a major in pre-occupational therapy, the other pursuing teacher certification in elementary education but not teacher certification in special education. Demographic identities by race or ethnicity included two black (one African American, and one European), one Hispanic, and one Native American. Three of the students disclosed disabilities, providing letters of accommodation from the Office of Academic Accessibility. Only two participants were males, the other 18 female.

Gains in Self-Efficacy

To investigate relationships between gains in self-efficacy and the quality of personal finance lesson design, researchers first computed gains in self-efficacy for overall self-efficacy for multi-tiered instruction and for five subdomains of multi-tiered self-efficacy. Students responded to the self-efficacy surveys at the beginning of the semester and then again at the end of the semester.

Gains in Mean Self-Efficacy

After each survey, researchers computed means for each student's responses in each subdomain and overall self-efficacy, and then the mean scores for all participants in each subdomain and overall self-efficacy. Students made gains in mean scores for overall self-efficacy for multi-tiered instruction (.72, *STD* .49), and in five subdomains: Engagement (.47, *STD* .72), Data-Driven Decision-Making (.64, *STD* .57), Collaboration (.59, *STD* .74), Evidence-Based Solutions (.73, *STD* .65), and Monitoring Interventions (.90, *STD* .61). In other words, students gained in all five subdomains of self-efficacy as well as overall self-efficacy for multi-tiered instruction.

Significant Self-Efficacy Differences

To test for significance of those gains in self-efficacy, the researchers computed two-tailed paired t-tests comparing post-semester self-efficacy scores with pre-semester self-efficacy scores. In the two-tailed paired t-test, results showed the differences were significant for overall self-efficacy and for four subdomains. In self-efficacy for overall multi-tiered instruction, students gained significantly across the semester, t=3.36, p<.01. Student gains in Engagement failed to reach statistical significance, t=1.80, p=.08. In the other subdomains, student gains did reach statistical significance: in self-efficacy for Data-Driven Decision-Making, t=3.10, p<.01; in self-efficacy for Collaboration, t=2.14, p<.05; in self-efficacy for finding and evaluating Evidence-Based Solutions, t=3.04, p<.01; and in self-efficacy for monitoring interventions, t=3.54, p<.01.

Effectiveness of Self-Efficacy Gains

The researchers then used Cohen's d to compute the magnitude of those gains, which reached significance in the two-tailed paired t scores. The researchers used the formula, Cohen's $d = (M_2 - M_1)/STD_{pooled}$ where $STD_{pooled} = \sqrt{((STD1^2 + STD2^2)/2)}$. Cohen (1988) initially established threshold for statistical significance with effect sizes equal to or greater than .40 considered statistically significant and those equal to or greater than .80 considered large. After increased research into effective teaching practices, Fuchs et al. (2017) later revised the statistical significance of effect sizes within the field of education, establishing that effect sizes between .25 to .34 are small but effective, between .35 to .49 are moderate, and effect sizes equal to or greater than .50 is strong.

Using Cohen's *d* to compute the magnitude of that difference, results showed that students made gains with a very large effect in overall self-efficacy for multi-tiered instruction, Cohen's d= 1.12. In all four subdomains with statistically significant t-score gains, students gained self-efficacy with magnitude of very large effectiveness: in Data-Driven Decision-Making, Cohen's d=1.04; in Collaboration, Cohen's d=.71; in Evidence-Based Solutions, Cohen's d=1.18; and in Monitoring Interventions, Cohen's d=1.18 (Table 1).

Table 1

Mean Scores, Gains and Significance in Self-Efficacy Domains of Multi-Tiered Instruction

SELF-EFFICACY FACTOR	PRE-TEST MEAN (STD)	POST-TEST MEAN (STD)	GAINS MEAN (STD)	2-TAILED PAIRED *t*	COHEN'S *d*
ENGAGEMENT	2.41(.68)	2.87(.82)	2.87(.82)	1.80	---
DATA-DRIVEN DECISION-MAKING	1.90(.67)	2.61(.69)	2.61(.69)	3.10**	1.04
COLLABORATION	2.19(.85)	2.76(.75)	2.76(.75)	2.14*	.71
FINDING AND EVALUATING EVIDENCE-BASED SOLUTIONS	2.36(.70)	3.05(.65)	3.05(.65)	3.04**	1.02
MONITORING INTERVENTIONS	1.98(.82)	2.95(.82)	2.95(.82)	3.54**	1.18
OVERALL MULTI-TIERED INSTRUCTION	2.13(.69)	2.88(.65)	2.88(.65)	3.36**	1.12

* means p <.05 ** means p <.01 --Cohen's d only computed if t-score significant

Relationships between Self-Efficacy and Lesson Quality

At the end of the personal finance module, the pre-service teachers designed brief video lessons on narrow topics about personal finance. The course instructor scored each video lesson by a rubric honoring such components as alignment with standards, demonstration of evidence-based practices, active engagement, having students use interactive materials, and overall video quality. Students also earned points for participation and giving meaningful peer feedback.

Quality of Lessons about Personal Finance

The quality of lesson design about personal finance as scored by the rubric varied significantly. On a scale of 100 points, scores ranged from a low of 5 points to three high scores of 100 points each. For lesson quality, the students earned a mean score of 77.5 (*STD 22.97*).

Correlations between Self-Efficacy and Lesson Quality

Investigating the relationships between those self-efficacy gains and the quality of the pre-service teachers' personal finance lesson design, the researchers analyzed correlations between those factors. Using SPSS, the researcher computed linear correlation between each set of two factors using two-tailed significance of Pearson's *r*. As shown in Table 2, the researchers found six statistically significant interactions between lesson quality and post-self-efficacy scores in overall multi-tiered instruction and in four sub-domains. Results demonstrated positive correlations between the quality of lessons designed by the pre-service teachers and their

self-efficacy for engagement of learners (r=.574, p<.05), data-driven decision-making (r=.670, p<.01), finding and evaluating evidence-based practices (r=.582, p<.01), monitoring interventions (r=.547, p<.05), and overall self-efficacy for multi-tiered instructional practices (r=.656, p<.01).

Table 2: Correlations between Lesson Quality and Post- Self-Efficacy for Multi-Tiered Instruction

Correlations between Lesson Quality and Self-Efficacy for Multi-Tiered Instruction

POST- SELF-EFFICACY FACTORS	CORRELATION WITH LESSON QUALITY
ENGAGEMENT	.574*
DATA-DRIVEN DECISION-MAKING	.670**
COLLABORATION	.257
FINDING AND EVALUATING EVIDENCE-BASED SOLUTIONS	.582**
MONITORING INTERVENTIONS	.547*
OVERALL MULTI-TIERED INSTRUCTION	.656**

* means p <.05 ** means p <.01

Furthermore, as shown in Table 3, results showed statistically significant positive correlations between lesson quality and gains in self-efficacy for Data-Driven Decision-Making (r=.507, p<.05). In other words, self-efficacy for components of multi-tiered teaching practices interacted significantly with the quality of teaching demonstrated through design of an instructional video on a narrow topic of personal finance.

Table 3: Correlations between Lesson Quality and Self-Efficacy Gains in Multi-Tiered Instruction

Correlations between Lesson Quality and Self-Efficacy Gains in Multi-Tiered Instruction

GAINS IN SELF-EFFICACY FACTORS	CORRELATION WITH LESSON QUALITY
ENGAGEMENT	.230
DATA-DRIVEN DECISION-MAKING	.507*
COLLABORATION	-.156
FINDING AND EVALUATING EVIDENCE-BASED SOLUTIONS	.187
MONITORING INTERVENTIONS	.450
OVERALL MULTI-TIERED INSTRUCTION	.399

* means p <.05 ** means p <.01

Effectiveness in Building Personal Finance Knowledge Test Function

For this portion of the study, researchers deleted data for one participant who failed to complete the pre-assessment in timely fashion. Then to ensure this researcher-designed test worked as intended, the faculty member completed item analysis. For test items upon which less than 55% of students scored correctly, the faculty set the discrimination index threshold of .30. A discrimination index of .30 or higher indicates a strong correlation between correct response on the item and strong performance on the test. On each of the 18 test items, students either performed well or the discrimination index was strong.

Gains in Knowledge of Personal Finance

Over the personal finance training module, the nineteen students made gains in personal finance knowledge as assessed on the test, with the same test used for both pre-assessment and post-assessment. On a scale of 100, student performance on the pre-assessment ranged from a low score of 14.29 to a high score of 88.57, with a mean pre-assessment score of 45.11 (*STD 18.16*). Student performance on that same test completed after the training ranged from a low of 17.5 to a high score of 100. Results showed a mean post-assessment score of 68.54 (*STD 22.70*).

The researchers did not anticipate factoring for the students' time to complete the test. Though students could take up to 120 minutes to complete the assessment, all took less than that allotted time. Students who took 20 minutes or less on the post-assessment demonstrated no gains or even lower scores on the post-test than on the pre-test. The student with the lowest score took ten minutes to complete the assessment while the student with the highest score completed the post-test in 35 minutes. The student with the highest gains of 62.86 points between pre-test and post-test took 39 minutes to complete the post-test.

To test for significance of gains between the pre-test and post-test, the researchers computed two-tailed paired t-tests comparing post-test performance with pre-test performance. In the two-tailed paired t-test, results showed the difference between the two means as statistically significant, $t\ (18) = 3.51\ (p<.01)$. In other words, results showed a statistically significant difference from pre-test to post-test of personal finance knowledge. The students did make significant gains in personal finance knowledge in response to the personal finance training module.

Gains Effectiveness

To investigate the second research question about effectiveness for outcomes of personal finance knowledge, the researchers then used Cohen's *d* to compute the magnitude of those gains. Results showed that students made gains with a very large effect, Cohen's $d= 1.14$. In other words, the instructional unit about personal finance resulted in a very large statistically significant gain from pre-test to post-test of knowledge about personal finance (Cohen's $d= 1.14$). Therefore, the instructional model about personal finance was effective in building knowledge and skills of personal finance.

Qualitative Analysis

Three researchers qualitatively analyzed the students' short written reflections about their experiences during the budget game, using color-coding to identify common themes. The researchers then discussed differences in theme coding, coming to shared agreement of three common themes and text fitting each theme. Student responses fit three themes: seven students expressing that Decisions Impact Finances, seven students discussing awareness of Needs versus Wants, and twelve students discussing the importance of Planning for the Unpredictable.

Decisions Impact Finances

Students expressed awareness of just how much decisions influence finances. Two students considered choices of positions, living arrangements, transportation, and benefits when reflecting on final budget outcomes. One student considered, "I wondered how my results would be different if I changed by initial responses...." Another shared that they "...learned that having someone to help you split the costs of everyday life is extremely important and can ultimately help you save money...." One shared realization of "...how many withdrawals there are and how spare the deposits are." Some students reflected on helpfulness of budgets. One expressed, "...having a budget makes spending more for important events like a friend's wedding more possible." Similarly, one student said that engaging in this budget game "...allowed me to have a more realistic experience with how personal finance looks for a teacher. There were many expenses that I never would have thought about. I didn't think about a friend's wedding or a friend having a baby. This game also made me think about planning for my personal financial future as a teacher. I never thought about the possibility of creating a *GoFundMe* for classroom expenses." These seven students expressed that decisions about living and benefits before starting the game, and decisions throughout the game influenced their finances.

Needs versus Wants

Through reflections, some students processed how they reasoned about differences between their needs and wants. One student shared that "it is important to remember that there are things that you may want to buy and there are things that you will need to buy as well. Therefore, it is important to spend wisely." Another pointed out, "...When it came time for big expenses like the wedding, I had to make a choice about whether attending would be a wise decision for me even though I really wanted to go. I also noticed how that having insurance for almost everything benefited me in the long haul, especially during unexpected medical expenses." Seven students discussed such wrestling with the differences between their needs and wants.

Planning for Unpredictable

Finally, twelve student reflections on the game supported the importance of planning for the unpredictable. One said, "There were many times during the game that I would save money, and then something unexpected would happen." Another expressed, "While playing the game I was at some highs and then with the flip of my coin, I would lose money... really shows how unpredictable life is." Some were specific

about learning, such as, "My eyes were opened to the unexpected medical bills that could happen and the life events that can take place." One expressed, "I saw in a tangible way how you can plan for consistent expenses like housing, and transportation. However, at the same time, this game also brought to life how various aspects of life are unplanned..." One student concluded, "If you are intentional about saving money, you will have a cushion to fall back on when you get in a rough spot or something pops up that you need to pay for." Through the forty simulated events of the budget game, twelve of the students appreciated how quickly circumstances could change around them.

DISCUSSION

Relationships between Post- Self-Efficacy and Lesson Quality

The strong correlations between factors of self-efficacy and quality of the lesson about personal finance fit previous research showing the importance of self-efficacy to teaching success (Burchard & Vargas, 2020; Barnes & Burchard, 2016; Erdem & Demirel, 2007; Woolfolk-Hoy, 2009). For example, when a teacher feels strong self-efficacy in finding and evaluating evidence-based teaching practices, that self-efficacy supports demonstration of such evidence-based practices. Likewise, when a teacher feels strong self-efficacy in engaging students, actively engaging one's students comes more naturally. Such results show the importance of teacher educators to build specific content knowledge and skills with concurrent support to the growth of self-efficacy of pre-service teachers, achievable through field practice, and timely and specific feedback, both corrective and positive.

Relationships between Self-Efficacy Gains and Lesson Quality

The strong correlation between lesson quality and gains in self-efficacy for data-driven decision-making requires a bit more consideration. First, correlations do not indicate causation or effect. The faculty assessed gains in self-efficacy from beginning of the semester to end of the semester.

All students pursuing teacher certification in this course enrolled in a concurrent math methods course as well as primary grades field experience. For the math methods course, each student completed a semester-long data-driven decision-making project gathering screening data on a small group of struggling students, designing an intervention, and monitoring progress in response to the intervention. The impressive group gains in self-efficacy for data-driven decision-making likely occurred in response to that curricular requirement. That explicit training and practice in how to design interventions to meet specific math learning needs prepared students well for the requirement in this course to demonstrate one lesson on a narrow slice of a personal finance topic that work for interventions in a personal finance course. Such results certainly show the value of integrating teacher preparation courses and field experiences.

Effectiveness of Training for Outcomes of Growth in Personal Finance Knowledge

Most impressive, the students in this course responded effectively to training about personal finance. Students' scored significantly higher on mean post-test scores of personal finance knowledge than on the mean

pre-test scores of personal finance knowledge, t (18) = 3.51 (p<.01). Results also showed strong magnitude in the difference from pre-test to post-test, with a very large effect size, Cohen's d= 1.14. Given the small sample size, this is unexpected, and encouraging for the value of such a unit to prepare pre-service special educators to teach or co-teach personal finance content.

Themes in Student Reflections

After coding the meaning of the student reflections on their experiences during the budget game, researchers identified three shared themes. Students expressed realization that decisions impact finances, that a simulation game helped them consider needs versus wants, and about planning for the unpredictable.

Limitations

The researchers note four limitations. First researchers investigated these questions at only one university, with a small sample size of only 20 participants. This limits broad conclusions even for implications for other pre-service teachers. In this study, one researcher was the instructor of the course. While pressures to participate were mitigated by study design, an instructor interpreting effectiveness results of their own teaching introduces risk for bias in interpretation of those results. Finally, because of a pandemic, participant self-ratings of efficacy were not paired with measures of effective co-teaching performance in high school classrooms, only with measures of university classroom performance. For all of those reasons, the researchers caution against broad conclusions of study results.

Next Steps

As the demand for financial literacy grows, so will the need for pre-service teacher training in personal finance, including for pre-service special educators who need skills to co-teach such content. The researchers intend to use this study to guide further curriculum improvements, perhaps increasing time devoted to this module when the university calendar returns to post-pandemic routines, and then investigating effectiveness of the refined instructional module for future student cohorts. For teacher preparation, the researchers recommend scaling such a study to larger sample size.

Furthermore, future research could investigate if gains in financial literacy by high school students are related to training of their high school teachers in such content. Of particular interest, future research might investigate if such gains in financial literacy by high school students are related to content preparation of co-teaching special educators.

CONCLUSION

High school courses in personal finances are one option for supporting transition goals toward independent living for many students in special education. Special educators need preparation to co-teach such content. This study showed interactions between teachers' self-efficacy for multi-tiered instruction and the quality of teaching personal finance content. Furthermore, this study demonstrated that an instructional module to train pre-service special educators in introductory concepts of personal finance worked with strong effectiveness (Cohen's d= 1.14). Qualitative analysis revealed three common themes of "wants versus needs", "decisions impact finances", and "planning for the unpredictable". Because researchers conducted this study in one university with a small sample size, they shy from broad conclusions, but acknowledge the importance of such results to guide future scholarship of teaching and learning about preparation to teach personal finance.

REFERENCES

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Barczack, M. A. (2019). Simulated and community-based instruction: Teaching students with intellectual and developmental disabilities to make financial transactions. *Teaching Exceptional Children, 51*(4), 313-321. https://journals.sagepub.com/doi/10.1177/0040059919826035

Barnes, S. K., & Burchard, M. S. (2010). *Developing a scale to assess self-efficacy for Response to Intervention practices in schools.* [Paper presentation 2009 Northeastern Educational Research Association]. Digital Commons @ Uconn at http://digitalcommons.uconn.edu/nera@2009/14/

Barnes, S. K., & Burchard, M. S. (2011). Quality and utility of the multi-tiered instruction self-efficacy scale. *Research and Practice in Assessment, 6*, 22-42. https://mosaic.messiah.edu/edu–ed/26/

Barnes, S. K., & Burchard, M. S. (2016). Utility of the Multi-tiered Instruction Self-Efficacy Scale in assessing needs and short-term gains of pre-service teachers for multi-tiered instruction. *Teacher Education and Practice, 29*(1).

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate.* Carnegie Foundation for the Advancement of Teaching: Princeton University Press.

Burchard, M., Dormer, J., & Fisler, J. (2017). Interactions between teachers' attribution for student learning and implementation of evidence-based practices. *AASA Journal of Scholarship & Practice, 14*(3), 28–44.

Burchard, M. S., & Myers, S. (2019). Early information literacy experience matters to self-efficacy and performance outcomes in teacher education. *Journal of College Reading and Learning, 49*(2). https://doi.org/10.1080/10790195.2019.1582372

Burchard, M. S., & Vargas, K. H. (2020). Interactions of IEP quality and intervention self-efficacy for pre-service teachers. *Pennsylvania Teacher Educator.* https://www.pac-te.org/

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Council for Exceptional Children, CEC. (2015). Initial special education preparation standards. *What every special educator must know: Professional ethics and standards.* https://exceptionalchildren.org

Erdem, E., & Demirel, O. (2007). Teacher self-efficacy belief. *Social Behavior and Personality, 35*(5), 573-586. https://doi.org/10.224/sbp.2007.35.5.573

Friend, M. (2016). Welcome to co-teaching 2.0. *Educational Leadership, 73*(4), 16-22.

Fuchs, L. S., Fuchs, D., & Malone, A. S. (2017). The taxonomy of intervention intensity. *Teaching Exceptional Children, 50*(1), 35-43. http://doi.org/:10.1177/0040059917703962

Henning, M. B., & Johnston-Rodriguez, S. (2018). Evaluating financial literacy curriculum for young adults with special needs: A review of content, universal design for learning, and culturally responsive curriculum principles. *Citizenship, Social and Economic Education, 17*(2), 118-135. https://doi.org/10.1177/2047173418789593

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004). https://sites.ed.gov/idea/statuteregulations/

Messiah University. (2020). Messiah at a Glance. https://www.messiah.edu

FINRA Investor Education Foundation. (2018). *National Financial Capability Study.* https://www.usfinancialcapability.org

Pennsylvania Department of Education. (2019, July). *Certification preparation program guidelines for special education grades PK-12.* https://www.education.pa.gov/Educators/Certification

Public School Code of 1949- Omnibus Amendments, P.L. 662, No. 91, Pennsylvania. https://www.legis.state.pa.us

Rodgers, W., & Weiss, M. P. (2019). Specially designed instruction in co-taught secondary mathematics courses. *Teaching Exceptional Children, 51*, 276-285. https://doi.org/10.1177/0040059919826546

Rosenbaum, E. (2020, Feb 5). *How each US state is shaping the personal finance IQ of its students.* CNBC. https://www.cnbc.com/2020/02/05/how-each-us-state-is-shaping-the-personal-finance-iq-of-students.html

Stoddard, C., Urban, C., & Schmeiser, M. D. (2018, November). College financing choices and academic performance. *Journal of Consumer Affairs, 52*(3), 540-561. https://doi.org/10.1111/joca.12175

Weiss, M. P., & Glaser, H. (2019). Instruction in co-teaching in the age of Endrew F. *Behavior Modification. Advance online publication.* https://doi.org/10.1177/0145445519836071

Woolfolk-Hoy, A., Hoy, W. K., & Davis, H. A. (2009). Teachers' self-efficacy beliefs. In K. R. Wenzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 627–653). Routledge/Taylor & Francis Group.

ABOUT THE AUTHORS

Melinda S. Burchard, Ph.D. is an Associate Professor of Special Education at Messiah University. Her research interests include self-efficacy and intervention skill development of pre-service teachers, especially pre-service special educators.
mburchard@messiah.edu

Alyssa Amels is a Second Grade Teacher in New Jersey, holding dual teaching certification in elementary grades and special education. Alyssa is passionate about STEM education in authentic and virtual experiences.
aamels21@gmail.com

Karissa Hagar is a Kindergarten Teacher in New York, holding dual teaching certification in elementary grades and special education. Karissa is passionate about serving students with disabilities in both the general and special education settings.
karissahagar@gmail.com

Carol Buckley, Ed. D. is an Associate Professor of Math Education at Messiah University. Research interests include strategies for developing numeracy, and supporting parents understanding of current Math instruction so they may help their child become successful in Mathematics.
cbuckley@messiah.edu

EXAMINING THE AFFECTIVE DIMENSIONS OF THE SOPHOMORE YEAR EXPERIENCE USING SITUATIONAL JUDGMENT TESTS

Amml Hussein, Ed.D.

PENNSYLVANIA EDUCATIONAL LEADERSHIP | VOLUME 41, NUMBER 1

ABSTRACT

This study explored the relationship between social and emotional learning variables on student success and retention during the sophomore year of the undergraduate experience. An examination of the affective levels of second- and third-year students commenced as measured by responses to a situational judgment test. A validated situational judgment instrument (SJT-AG) (Westring et al., 2009) was administered to a diverse sample of second- and third-year students at a four year state insitution on the east coast of the USA. The situational judgement test (SJT) assessed participants along two dimensions: behavioral responses and achievement goal orientation. The following research questions were addressed: RQ1: What affective factors/ psychosocial factors impact student outcomes of second year undergraduate students? RQ2: How do higher scores on the SJT (affective variables) affect student outcomes in terms of course performance and retention? The study concluded and offered recommendations for student support services and retention initiatives.

Keywords: situational judgement tests, affective variables, collegiate success, retention

INTRODUCTION

The sophomore year, widely referred to as the "sophomore slump," is a stage of the undergraduate experience characterized by reduced motivation, increased departures, and low GPA performance. For decades, scholars have explored ways to improve institutional academic support services to help college students (Tinto, 1975, 1982; Schreiner et al., 2012; Zaitseva et al., 2013). Researchers have identified the sophomore year as an understudied stage of the undergraduate journey (Boivin et al., 2000; Schreiner et al., 2000; Graunke & Woosley, 2005) whereby students face distinct challenges (Schreiner et al., 2012). The sophomore year warrants further investigation due to its unique complexities (Hunter et al., 2010; Yorke et al., 2015; Webb & Cotton, 2019). Increased pressure for students to declare a major and career path, while course work progressively becomes more challenging, coupled with a decreased level of institutional supports makes the sophomore year especially difficult for students to navigate (Schreiner et al., 2012; Milsom & Yorke, 2015). Meanwhile, colleges are interested in methods to better meet student needs and prevent departures while supporting students through this senstive stage. Prior studies also point to the important role that student support services play with respect to preventing college departures (Rodríguez-Muñiz et al., 2019).

There is an evident gap in the literature related to the sophomore year population and researchers have dubbed sophomores as "invisible students" (Schreiner & Pattengale, 2000; Zaitseva et al., 2013). The existing literature is predominantly descriptive in nature, providing descriptions of the problem (Schreiner et al., 2012) or calling for enhanced institutional responses (Tobolowsky & Cox, 2012; Hunter et al., 2010). Since the literature is two-fold: partly descriptive and partly calling for institutional response, this study sought to uncover the unique challenges faced by second year students using an exploratory design collecting data from second and third year students. This author hypothesized that if insitutions gain access to psycho-social and affective domains of their enrolled students, this information may be combined with machine learning and data mining techniques to enhance insitutional responsiveness to the sophomore year population in particular.

Research Problem

One primary theme in college marketing strategies is that institutions seek a broader more complex student as evidenced by the admission application pamphlets for prospective students (Oswald et al., 2004). Just as colleges have identified their desire for well-rounded students, the selection process should also include broader measures and more complex combinations of individual abilities to predict performance and academic success (Oswald et al., 2004). Developing linear regression models that examine students' achievement goal orientation through an SJT instrument might help address this problem and promote a balance between cognitive predictors and more complex combinations of individual characteristics, also known as affective variables, that may predict performance, persistence, and academic success. To better understand the second year student experience at the host institution, this researcher examined the implications for using situational judgment inventories to predict student outcomes. SJTs are easily administered electronically using computers to help improve outcomes and are easily scored electronically. These benefits are attractive to instituutions who are looking for ways to improve student outcomes, improve retention rates, and enhance institutional response to students.

LITERATURE REVIEW

Prior research suggests that goal orientation may help researchers better understand student achievement and retention (Westring et al., 2009). Mastery goal orientation involves mastering a subject for the sake of learning the discipline; performance orientation is related to students wanting to outperform others and seeking the approval of their assessors and peers (Senko & Harackiewics, 2002). Midgley et al. (2001) demonstated a positive association between performance orientation and course grades and researchers have indicated that better grades may be due to shallow study behaviors and memorization.

Mastery Goal Orientation

Prior studies link mastery goal orientation with a more enjoyable student experience (Midgley et al., 2001) and stronger grades in courses that assess complex challenging material (Grant & Dweck, 2003). Stewart and Darwent (2014) found that mastery goal orientation was associated with poorer academic performance during second year cohorts.

This researcher's point of intrigue was from the vantage point of examining goal orientation systematically through scores on a situational judgement tests (SJTs), while drawing upon instiutional archival records. SJTs have the potential to measure with precision each study participant's achievement goal orientation based upon their responses to situational scenarios that they are likely to encounter as undergraduate students. Students were asked to rate how much they agree with each response item using a 4 point Likert scale. This researcher hypothesized that using an SJT would help create a formula for institutions to utilize in order to interpret their archival records using machine learning and data mining techniques. This method may translate into institutional opportunities for targetted support and enhance the sophomore student experience by optimizing student support and retention interventions using population segmentation procedures.

Utility of Situational Judgements

Situational judgment tests (SJTs) have been utilized widely in the occupational/employment arena to predict employee retention (O'Connell et al., 2007). Komarraju et al. (2013) were interested in exploring affective dimensions of students to gain an understanding of their goal orientation, goal commitment, and situational decision making. A handful of studies were found in the extant literature that explored SJTs in high stakes settings including medical schools, dental school, and military program selection (Guenole et al., 2017; Sharma et al., 2013; Patterson et al., 2009, 2016; Ashworth et al., 2014; Sartania et al., 2014; Lievens et al., 2005; Whetzel & McDaniel, 2009; de Leng et al., 2018). This study sought to use the SJT as a metric to quantify sophomore and junior students' achievement goal orientation across disciplines at a 4 year state institution, which was another evident gap in the literature as studies that examined the sophomore year using SJTs were absent from the extant literature.

METHODS

A validated situational judgment test (SJT-AG) (Westring et al., 2009) was administered to a sample of second and third year students during the fall 2019 semester. All SJT response items had uniqueness of .30 or above (). (Westring et al., 2009). Moreover, Confirmatory Factor Analysis (CFA) was excellent fit to the data. The SJT-AG has a Tucker Lewis Index=.98. The Tucker-Lewis index is used to resolve issues pertaining to negative bias. A full IRB proposal review was completed and approved for this study at the host institution.

Research Questions. This study addressed the following research questions:
RQ1: What affective factors/psychosocial factors impact student outcomes of second year undergraduate students?
RQ2: How do higher scores on the SJT (affective variables) affect student outcomes in terms of course performance and retention?

This study measured affective levels of second and third year students as quantified by the SJT-AG (Westring et al., 2009), a validated instrument with high levels of internal and external validity.The SJT-AG instrument outlined 8 scenarios that students are likely to encounter during their undergraduate college experience. Knowledge instructions accompanied each situational judgement test. Each item measured goal orientation upon three dimensions: mastery, performance approach, and performance avoidance orientations. All responses were entered by the respondents as values on a 4-point Likert scale. Participants were provided with a test sheet and scantron, and were instructed to use their student identification number as their only identifier.

Recruitment

This researcher recruited students through a convenience, purposive sample. Ten minute presentations about the research study were given to several undergraduate courses across disciplines at the host institution targeting second and third year students.

Selection Criteria

Second year and third year students were recruited for the purpose of examining the sophomore student experience and better understand problems with retention. Using a second year sample provided this researcher with first year academic outcome information (a minimum of two semesters of course grades and cumulative GPA) were used as a mediating variable to interpret results.

Exclusion Criteria

First year students were excluded from the sample due to several confounding variables including adjustment issues, mental health confounders, campus climate adjustment, distance from family/support systems, and rigor of collegiate course work compared to high school. Therefore, this researcher sought to work with second year students who are more acclimated to the college environment. Moreover, scholars

have determined that personality traits such as motivation are fairly stable during the second year stage of undergraduate studies (Delaney et al., 2013; Cobb-Clark & Schurer, 2012; Rivkin et al., 2005). Fourth year students were excluded from the study.

The Sample

Table 1 represents a profile of the phase two sample by department. The table demonstrates that close to half of the sample were majoring in the sciences, while the other half were liberal arts and social science majors.

Measures Taken for the Protection of Participants

This researcher took several steps to protect the privacy of student subjects. Student identification numbers were the only identifier used on the survey and demographic forms. All data was stored in a password encrypted computer and data was stored in a locked cabinet with access only to this researcher.

RESULTS

Affective variables including achievement goal orientation may provide educators with the utility to predict student success outcomes including college cumulative GPA and retention (Plominski & Burns, 2018; Lee & Leonard, 2009; Richardson et al., 2012; Webb & Cotton, 2019; McAbee & Oswald, 2013; Oswald et al., 2004). Affective variables are not traditionally assessed or measured by institutions during the second year of the undergraduate years. Students at this stage of their coursework navigate a slew of situational and emotional circumstances that may implicate their student success outcomes.

Participant Demographics

Table 2 in the appendix provides summary statistics for the each of the subject variables. The sample mean scores for the Fall 2019 GPA and the HSGPA were 3.12 and 3.31, while the sample means scores for NSATERW and NSATM were 524 and 499. Table 3 gives the frequency counts and relative frequencies for RetSp2020, the students who were retained for the Spring 2020 semester. Table 3 shows that 90% (88) of the students in the study sample were retained for the Spring 2020 semester. Thus, 10 students departed following the fall 2019 semester, and 88 students were retained through spring 2020.

Descriptive Statistics

Descriptive statistics were utilized to provide a profile of the sample (See Table 2 for Summary Statistics). The performance variables used in this study include high school GPA, SAT scores, Fall 2019 GPA. The affective variables measured in this study include the following goal orientations: mastery goal orientation (C1), Performance approach goal orientation (C2), and Performance avoidance goal orientation (C3).

Statistical Analyses

The results of this study suggested a statistically significant negative correlation between SJT score in the performance approach category (C-2) and fall 2019 GPA. There were no statistically significant correlations

found for mastery approach category (C-1) or performance avoidance (C-3) and fall 2019 GPA. C2 was negatively correlated with Fall 2019 GPA (-.239*). Meanwhile, C1 had no statistically significant associations with either fall 2019 GPA or retention (See Table 4). Additional statistical analyses were conducted to determine the strength and direction of relationship between performance measures: college GPA and affective variables including goal orientation.

Correlations

There were no statistically significant findings related to mastery goal orientation. The performance approach variable (C-2) had a statistically negative association with Fall 2019 GPA (Table 4), and there were no significant associations between any of the SJT constructs and retention (Table 4). All three constructs were significantly correlated with one another, with a significant positive association between C1 – Mastery and C2 – Approach ($r = .61$, $p < .01$), a significant negative association between C1 – Mastery and C3 – Avoidance ($r = -.55$, $p < .01$), and a significant negative association between C2 – Approach and C3 – Avoidance ($r = -.52$, $p < .01$).

Table 5 shows the point biserial correlations between survey factors and fall 2019 retention. Point-biserial correlation coefficients were calculated for each of the constructs and retention because retention is a binary variable (i.e. it has two options, 1 for students being retained, 0 for students who are not retained the following semester), and the survey factors were interval-ratio variables. The results of the point biserial correlation coefficients indicate no statistically significant relationships between any of the survey factors and fall 2019 retention.

RESEARCH QUESTIONS ADDRESSED

Research question 1: What affective factors/psychosocial factors impact student outcomes of second year undergraduate students? To address research question 1, regression models between the independent variables and the two dependent variables, fall 2019 GPA and spring 2020 retention were investigated. For fall 2019 GPA, stepwise regression using the backward elimination method was used to find the model of best fit. Stepwise regression was selected because it is appropriate for exploratory models, and backward elimination was used because it minimizes suppressor effects better than other stepwise procedures (Field, 2013). Suppressor effects occur when a predictor has a significant effect but only when another variable is held constant.

Backwards elimination is a procedure that begins with all the study variables in the model, then variables are removed one by one, and the model is retested at each stage. This process was repeated with each of the variables until a final regression model was produced that had maximum statistical fit. The independent variables were entered into the model including mastery goal orientation, approach goal orientation, avoidance goal orientation, high school GPA, and math and verbal SAT scores. For spring 2020 retention, a logistic regression model was explored with the same independent variables as were entered into the initial regression model for fall 2019 GPA. A logistic regression model was used for spring 2020 retention because spring 2020 retention is a nominal variable and regular multiple regression is only appropriate for dependent variables that are interval/ratio.

Table 6 provides the results of the regression model for Fall 2019 GPA. The regression model was highly significant ($F(3, 73) = 5.89$, $p < .001$) and had two significant predictors, C2 – Approach ($B = - 0.03$, $p < .05$) and HSGPA ($B = 0.41$, $p < .05$). The negative coefficient for C2 – Approach indicates a negative relationship between C2 – Approach and Fall 2019 GPA, that is, as the score for C2 – Approach increases, the Fall 2019 GPA decreases. More specifically, as the C2 – Approach score increases by one unit, Fall 2019 GPA decreases by 0.03 if the effects of all other predictors are held constant.

The regression model from Table 6 can be summarized by the following formula: Fall 2019 GPA = 1.89 - 0.03 (C2 - Approach) + 0.41 (HSGPA). This formula provides a model for predicting GPA from the C2 – Approach scores and HSGPA. The r2 value for this model was .20, indicating that 20 % of the variability in Fall 2019 GPA was accounted for by C2 – Approach and HSGPA.

Affective Factors

The results of the logistic regression analysis for the dependent variable RetSp2020 are given in Table 7. The regression model was highly significant ($2(6) = 24.93$, $p < .001$) and had three significant predictors, C3 – Avoidance ($B = 0.32$, $p < .05$), NSATERW ($B = 0.02$, $p < .05$), and NSATM ($B = -0.03$, $p < .05$).

To interpret the coefficients of the predictors, this researcher examined the column labeled Exp(B). Exp(B) is the odds ratio, which is an indicator of the change in odds of the occurrence of the dependent variable resulting from a change in the predictor. Predictors with an odds ratio greater than one mean that as the predictor variable increases, so do the odds of the dependent variable occurring. Alternatively, for predictors with an odds ratio less than one, as the predictor variable increases, the odds of the dependent variable occurring decrease (Field, 2013). Thus, for the model in Table 6, the odds ratio for C3 – Avoidance is 1.38 indicating that as C3 – Avoidance increases, the odds of being retained also increases. However, the odds ratio for NSATM was -0.03, indicating that as the value of NSATM increases, the odds of being retained decrease. This demonstrates a problem with relying on SAT scores, as the higher the score, they appear to predict departures within this sample.

One of the measures used to assess the model fit for multiple regression models is r2, the proportion of the variance in the dependent variable accounted for by the independent variables in the model. There is no direct analogue to r2 for logistic regression, however, a number of r – statistics have been proposed to indicate the strength of model fit for logistic regression. One of the more commonly cited R – statistics, and one that SPSS reports, is Nagelkerke's r2N. For the logistic regression model in Table 7, SPSS reported a Nagelkerke r2N value of .50, indicating a fairly good model fit.

Research Question 2: How do higher scores on the SJT (affective variables) affect student outcomes in terms of course performance and retention? The results of the regression analyses were used to address Research Question 2. Specifically, higher scores on C2 – Approach have a negative effect on Fall 2019 grade point averages, while neither C1 – Mastery nor C3 – Avoidance have any effect on Fall 2019 GPA. Similarly, the results of the regression analyses suggest a positive relationship between performance avoidance (C-3) construct and Spring 2020 retention. Both mastery goal orientation (C-1) and Performance approach (C-2) had no significant relationship with Spring 2020 retention.

Retention

Performance avoidance (C-3) was positively correlated with spring 2020 retention. Mastery orientation (C-1) and Performance approach (C-2) had no significant relationships with spring 2020 retention. (See Table 6).

To summarize, the results suggest two important statistically significant relationships. A statistically negative relationship was found between performance approach (C2) and fall 2019 GPA, and a significant positive relationship was found between performance avoidance (C3) and spring 2020 retention.

The following regression model was produced: Fall 2019 GPA= intercept + ß1 (C2-Performance Approach) + ß2 (HSGPA).

Coefficients

A Pearson's correlation co-efficient was calculated for each of the construct scores (C1-Mastery, C2-Performance Approach, and C3-Performance Avoidance (See Table 3). Pearson's correlation coefficient is symbolized as r, where r is between 0 and 1. Therefore, r helped this researcher describe the extent that the variables are related to one another. The result of this test found a negative relationship (as the value of x increases, the value of y (retention) decreases therefore, there was an inverse correlation discovered between performance approach and fall 2019 GPA. Additionally, a statistically positive correlation was found between performance avoidance goals and spring 2020 retention.

Next, the scores were converted into difference from the mean scores. Difference from the mean was calculated to determine whether the score is above or below the mean, and how far below or above it. When x and y scores are not related, the CP xy will be closer to 0 (Kleiss & Green, 2010). Then, this researcher calculated r2 (the co-efficient of determination) which "provides an analogous measure of the strength of the association between x and y." (Kleiss & Green, 2010, p. 391).

Statistical Significance of the Pearson r

A rejection region was located in the sampling distribution of R. Since R fell into the rejection region, this researcher rejected the null hypothesis and accepted the alternate hypothesis. This is because r falls into the rejection region and the probability of Rrobs is less than or equal to 0.05, the null is rejected, and the alternate hypothesis is accepted. Robs is statistically significant at the 0.05 level. Therefore, it was decided that performance approach and fall GPA variables are correlated in the population from which this sample was selected.

Assumptions. That SJT-AG item scores and GPA possess a bivariate normal distribution in the population from which the scores were sampled. The co-efficient of determination indicates that the SJT measures for performance approach (C-2) covaries with a negative relationship between approach and fall GPA. Another assumption was that students answered honestly and did not engage in faking their responses, meaning that students responded to each SJT item true to their own experience and not in the way that they think that the researcher would like them to respond.

Residuals. In order to calculate the error in prediction, this researcher calculated residuals, which are the measure of error in prediction. The residual was used as the basis for measuring the accuracy of the

prediction called the standard of error estimate (Kleiss et al., 2010). In order to develop a measure of error in prediction, the residual was squared. The standard of error of estimate was then calculated from the value of r.

Because the retention variable is a categorical variable (0 and 1), a Pearson Product was conducted. There was no statistically significant correlation found between affective constructs (C1, C2). A statistically signficant positive relationship was found between retention and C3-performance avoidance and spring 2020 retention. Finally, avoidance goal orientation (C3) covaried with retention with an r2 = .50 and a positive relationship = 0.32* (See Table 7). Retention was coded as a dichotomous variable 1=retained, 0=not retained in the spring semester.

Associations Between Constructs

Table 3 shows the associations between constructs. The results revealed statistically significant interactions between constructs: C1 and C2 (.61**). C2 and C3 had a negative association (-.52**) between the two constructs. Regression analyses permitted this researcher to explore relationships between performance measures and the three affective constructs. The findings suggest a positive relationship between high school GPA and college GPA (fall 2019). This researcher conducted backward elimination, in which all of the variables were entered, and they were removed one by one to see which of the constructs maintains a statistically significant relationship/impact. R2 was conducted. The r2 statistic (coefficient of determination) was selected because it helps gauge the closeness of the data fit to the regression line. The following constructs had the most significant associations: high school GPA, performance approach (negative association), and performance avoidance (positively correlated with retention).

Intercorrelations were conducted for each of the variables: dependent variable college GPA. The subject variables include C1, C2, C3, high school GPA, college GPA, SAT, and retention. The results of the summary statistics indicate that there are significant associations between college GPA and performance approach and performance avoidance constructs. There were no statistically significant associations found between fall 2019 GPA and mastery orientation (C1). The results also indicate no statistically significant associations between retention and C1, C2.

To summarize, the findings revealed three statistically significant associations (See Figure 1). A statistically negative association was found between performance approach and fall GPA. A positive relationship was found between performance avoidance (C3) and spring retention. Finally, a positive correlation with statistical signficance was found between high school GPA and fall GPA.

SUMMARY OF THE RESEARCH FINDINGS

This study found that academic, situational, and psychosocial facets collectively play a role in the overall experience of the sophomore year. Understanding students in the context of these three broad facets provides us with a multi-dimensional, holistic perspective of the sophomore year experience. This study gathered quantitative data that demonstrated a negative relationship between performance approach goal orientation and GPA performance for the sample under study. Additionally, a positive relationship was found between high school GPA and tertiary GPA which was consistent with prior studies, and the performance avoidance goal orientation was positively associated with retention. These findings suggest that affective

measures may be quantified using SJTs and moreover, the affective dimensions are useful measures to include in predictive modeling at the institutional level.

DISCUSSION

The results of this study are compelling: the utility of using SJT-AG as a method to quantify affective levels and goal orientation of undergraduate students translates into an early detection system that institutions may utilize to strategize institutional response to the sophmore year population and target appropriate student support services to intervene with precision. The low press context of situational judgement tests, its relatively low cost, and ease of administration make it an attractive method to quantify students' affective dimensions, which traditionally have been untapped due to the difficulty in measuring affective variables. Fortunately, the results of the regression analysis suggest that the scores on SJTs are predictive of GPA and retention in future semesters.

The implications of these findings for retention program administrators is quite significant because administrators may forecast student GPA for the next semester using the GPA regression model developed by this author. Coincidentally, institutions may target their retention initiatives more strategically and provide additional student support services to the students who score particularly high on the performance approach construct. Very few studies have used the SJT-AG specifically for predicting second year performance outcomes and retention, which emphasizes the novelty and uniqueness of this research study.

This study found that performance approach goals were negatively associated with GPA. The results of the backwards regression analyses revealed a negative relationship between performance approach goal orientation and fall 2019 GPA for the sample under study. This finding was inconsistent with Hsieh et al. (2007) findings which found mastery goal orientation and avoidance goals were the strongest predictors of college GPA, meanwhile, the instant study did not find any statistically significant relationships between mastery goal orientation and GPA. One explanation for the difference in findings could be the difference in the instrumentation. This study utilized a situational judgment test to measure goal orientation, which differs from the instrument used in Hsieh et al. (2007) study. Another potential explanation for the inconsistent finding may be that this researcher created specific parameters for the sample population and limited the study to second and third year students only. It is possible that morre advanced students such as seniors, or alternatively less advanced freshmen, may account for the observed differences in the associations between mastery approach and GPA in the Hsieh et al. (2007) study. Furthermore, the Hsieh et al. (2007) study explored a high rate of students on academic probation. The sample for the instant study included students with good academic standing.

Statistical Analysis

According to the regression model, students that scored high in the performance approach category would be more likely to have lower GPAs. One potential explanation for this finding is that due to an external locus of control, students inherently pay less attention to self-regulatory strategies and learning processes, and this would help explain why students who have high levels in the performance approach category would be more likely to attain lower GPAs. Conversely, if students score lower in the performance approach category of

the SJT-AG, they would be more likely to earn higher GPAs in the next sequential semester.

This study also found that performance avoidance orientation was positive associated with retention. Because there was a high retention rate for this particular sample these findings must be taken with caution. Furthermore, previous studies have not found statistically significant correlations between affective subject variables and retention (Nieves, 1991).

Lastly, the odds ration for NSATM the odds ratio was -0.03, indicating that as the value of NSATM increases, the odds of being retained decrease. This highlights a significant problem with relying on SAT scores, as the higher the score on the NSATM, appeared to predict departures in this study's sample. Thus, alternative evaluations for college admissions may be warranted. SJTs hold promise in the college admissions arena where they may be implemented to provide a fuller picture of each applicants' strenghts, especially non-cognitive skills that are not present anywhere else in admission applications.

Recommendations for the Application of the Current Study

This researcher offers several recommendations for implementing the findings of this research study into practice and policy which are set forth below.

Detection System. It is recommended that institutions extend these findings to improve our conceptions of the sophomore year stage. By combining cognitive predictors with affective measures such as the ones measured in the SJT-AG as alternative predictors in a compensatory model, institutions may devise an early detection system of students who may at risk for departure, students at risk for low GPA/performance. This would permit institutions to target student support services to prevent college departures. At the same time, institutions gain a method to identitfy their highest performing students for meritorious recognition.

Predictive Models. There were two regression equations that were produced from this study, one for GPA predictions which would provide useful information for institutions that want to identify students at risk for scoring in lower GPA. The GPA equation model has two major benefits: first, it has the capability to predict highly achieving students for the purpose of administering scholarships and academic achievement awards. Secondly, the GPA equation may identify students at risk for scoring lower GPAs in their subsequent semester. The second equation developed as a result of this study predicts retention and may be used to identify students at risk for departure. Institutional response may be improved by targeting student support services to students who score in the at risk category and efforts to engage students, mentor them, and retain them on track towards degree completion would become more focused.

Using the regression model and regression equation may aid institutions in identifying students who may be at risk for poor performance or departure. Furthermore, the regression model identified one factor (performance avoidance) that had a statistically positive association with retention. Plugging in the responses in both constructs from an SJT regression equation could assist institutions in identifying at risk students by utilizing predictive modeling that forecasts students GPA and retention. The scores derived from SJTs may provide institutions with a robust method of identifying at risk students for the fall semester of the sophomore year. Students who score high on the approach construct may be given an intervention and additional student support services to intervene with the goal of improving retention.

The results of this study demonstrated that performance approach goal orientation was associated

with college GPA. These findings are compelling and consistent with Ackerman et al. (2013) who used trait complexes including science/math, mastery/approach-achievement motivation, verbal/intellectual, avoidance in achievement contexts, and social/extroversion. The first three were associated with being positively related to academic performance and mastering domain knowledge (Ackerman et al., 2013). The other two proved to have a negative effect on academic performance and the "acquisition" of domain knowledge. This means that this researcher's findings are parallel with Ackerman et al. (2013) findings that found that avoidance in achievement contexts and social/extroversion were negatively associated with GPA, in the same way, this study found that performance approach goal orientation was negatively associated with fall 2019 GPA.

Institutional Practice and Policy. Lastly, it is recommended that institutions examine policies from the perspective of sophomores. Sophomore year students are often last in line in terms of course registration, access to housing, and financial aid (Schreiner et al., 2012). Academic advising should be driven by developing educational goals and life goals for students, with the aim of building pathways to success. Using affective instruments such as the SJT-AG will help institutions measure student thriving during key transitional stages of the undergraduate journey. Finally, institutions should keep the big picture in mind when designing campus programs for second year students. Institutions may plan to assess student strengths at the beginning of the undergraduate experience to establish a baseline. Then, affective measures such as achievement goal orientation may be re-assessed at the end of the freshman year, and each semester thereafter to monitor student development.

Student/Faculty Research Partnerships

Sophomores may benefit from opportunities to interact with faculty which may help encourage a greater sense of belonging within the larger organization of the university (Graunke & Woosley, 2005; Hunter et al., 2010). Prior studies recommend investing in faculty/student research partnerships during the sophomore year to promote student engagement and satisfaction (Wilson & Crowe, 2010). For example, Nagda et al. (1998) found that sophomore and African American students benefited the most from research partnerships, which was also positively associated with retention. This study recommends an increase in sophomore specific programming, which may be achieved by institutions connecting faculty to sophomore students which would translate into numerous benefits that extend to all grade levels and into the workplace and beyond (Kuh & Hu, 2001; Pascarella & Terenzini, 2005).

Communications Between Institutions and Sophomores

It is recommended that universities maintain open communication with the sophomore student population. Several scholars agree that there is a need for increasing communication between institution and the sophomore student population (Boivin et al., 2000; Cheng, 2004; Schreiner et al., 2012). Increased communication between the sophomore student body and the institution may be achieved through various social media outlets, newsletters, or emails to students. These communication channels promote increased communication between students and the university, which may improve students' feelings of belongingness and connectedness to the institution.

Increasing institutional communications directly to the sophomore population is also recommended

and enlisting student feedback and being attentive and responsive to their feedback may foster increased sense of belonging for the sophomore population. This may be achieved by developing a cost effective electronic channel for student feedback surveys. Collecting data on a course by course basis provides instructors with valuable information about the affective levels of their students and course activities may be tailored to meet the needs of the students enrolled in each course (Graunke & Woosley, 2005; Hunter et al., 2010). Investing in communication channels whereby students raise concerns to the institutions gain access to student voices ultimately enhancing student engagement and satisfaction.

Community Building

Another way to implement this study's findings would be building communities inside and outside the classroom. This may be achieved by providing faculty with training to cultivate collaborative classrooms to "capitalize on students' strengths and learning styles can ensure that classroom becomes a place where sophomores feel a sense of belonging and connection" (Schreiner et al., 2012, p. 129). Capara et al. (2011) suggest that there is a need for "changes in instructional experience." These changes may be achieved through deliberate adjustments in classroom environments and training faculty to recognize these opportunities is essential.

Prior studies suggest that sophomores had considerably lower levels of psychological wellbeing (as measured by dysthymic affect) in both honor's and non-honor's groups. This has been widely acknowledged in the research community as "the sophomore slump" where students struggle to define themselves professionally, academically and personally (Gump, 2007; Tower et al., 2015; Wang & Kennedy-Phillips, 2013). Moreover, prior research demonstrates a significantly positive relationship between internal locus of control and psychological wellbeing (and individuals reported a greater sense of control and healthier levels of happiness (April et al., 2012; Mobarakeh et al., 2015; Myers & Diener, 1995).With fewer supports available compared to the first year experience, students experience an abrupt transition from first year to the sophomore year. Furthermore, the extant literature suggests that services are offered to freshmen, juniors and seniors, and the sophomore year tends to receive the least amount of institutional support (Sanchez- Leguelinel, 2008; Schaller, 2005). This presents an opportunity to improve internal policies to enhance the sophomore year experience. These findings help institutions develop a bridge to buffer the stark gap between the first year experience and second year experience and promote a smoother transition.

Student Support Services

The use of machine learning and data mining techniques that combine institutional records (Rodríguez-Muñiz et al., 2019) with SJT scores using regression models may help institutions identify students that are at risk for departure or poor GPA performance and interpret their data into translatable reports and visualizations. Using this information, institutions will have the means to target specialized student support services for students who score in the at risk categories. These programs may focus on personal and professional obstacles that are experienced in the second year (Graunke & Woosley, 2005; Hunter et al., 2010; Sanchez-Leguelinel, 2008; Schaller, 2005; Wang & Kennedy-Phillips, 2013). Creating sophomore specific programming helps promote a campus ethos that emphasizes student engagement and student learning

(Schreiner et al., 2012). Furthermore, institutions may strike a healthy dialogue with the sophomore student population. Cheng (2004) indicates that "what connects students with the community is not just small circles of friends who share personal interests; it is also effective programming and organized social opportunities" (p. 228). Therefore, it is worthwhile for institutions to design programs that engage students and seek their feedback in order to enhance the sophomore year experience. Academic advisors may work on goal setting with students, especially students who score in the at risk category using the two regression model equations derived from this study.

FUTURE RESEARCH

Learning Communities

There is a need to pilot and explore the efficacy of learning communities during the sophomore year. Previous studies examined the concept of learning communities in the context of first year programs (Hotchkiss et al., 2006) and found that freshmen learning communities were associated with increasing student GPA grade levels by one letter grade which varied depending on race and ethnic backgrounds. Other findings mentioned the use of learning communities as a potential avenue for institutions to increase sophomore specific programs for student engagement (Schreiner et al., 2012). In addition, smaller courses and learning communities designed specifically for second year may foster a sense of community for sophomore students. Future studies may evaluate the effectiveness of adding a service learning component to existing courses (Schreiner et al., 2012). Lastly, differences between performance outcomes and racial/ethnic groups are recommended for future explorations.

College Admissions Practices

The use of SJTs in the undergraduate admissions process is another direction for future research. The SJT may help undergraduate institutions make selection decisions that consider both based upon both cognitive and affective variables of the applicants using compensatory models. Patterson et al. (2016) recommend developing selection tools that measure non-cognitive attributes of college applicants to predict collegiate success. SJTs have been used to identify whether candidates are likely to succeed in medical training and advance to become competent clinicians. Kreiter & Axelson, (2013) used SJTs to support medical school admission decisions. Kreiter & Axelson (2013) suggest that using well designed selection tools results in better outcomes. Admission practices have moved away from subjective measures like personal statements and recommendation letters and moved in the direction of using evidence based models of selection, which may include SJTs (Kreiter & Axelson, 2013).

Combining cognitive predictors like standardized tests, with non-cognitive variables, in a linear regression model permits institutions to select individuals with lower levels of cognitive abilities, but overall desirable college applicants that may persevere. An abridged SJT may be used to gather affective measures on all applicants. Also, the SJT may be used as a method to sort and prioritize applicants on waiting lists.

Finally, because of the concern with the adverse impact resulting from standardized cognitive ability and achievement tests, there is a need to measure mean differences across racial and gender subgroups. Historically, the SAT/ACT had notable differences in test performance scores across different racial

backgrounds (Culpepper & Davenport, 2009). Gaining insight into students' affective measures may help supplement disadvantaged students with their college admission applications and provide institutions with factors that are associated with predicting GPA and retention as established by the regression equation derived from this study. SJTs have practical utility for admissions personnel (Oswald et al., 2004) and may provide a fair and standardized method to score applicants on a broad range of prior social and educational experiences. This added component may supplement the existing online admissions application and provides the admission committee with a unique and personal cross section of each applicant. Furthermore, the SJT is easily scored, and these items may help predict or identify risk/vulnerability for college departure using regression models like the two developed in this study. This study found that performance avoidance goal orientation was positive correlated with student retention with the study sample. Institutions may integrate SJT scores to interpret their archival and institutional data using machine learning to enhance their retention interventions.

Limitations

This study was limited to a modest sample size. The findings are related to this specific sample. Therefore, future studies and replicatonis of this study are warranted on larger samples in order to extend the generalizability of these findings. Additionally, this study did not gather racial/ethnic/identity data, therefore, generalizations and differences across racial and ethnic identities cannot be gauged.

CONCLUSIONS

The sophomore year is a complex stage of the undergraduate experience. This study examined the sophomore year in a multi-dimensional fashion from the perspective of sophomore and junior students that successfully navigated the sophomore year. The findings suggest that students demonstrate a performance orientation that drives their study habits. Additionally, the study found that performance approach goal orientation is negatively associated with GPA performance. A positive relationship was found between performance avoidance and retention. Using the two regression equations, institutions may identify at risk students and target student support initiatives specifically to the identified students. SJTs may be a viable method that utlizes computers to help institutions interrpret archival records combined with students' current performance. The tests have applicability in admissions and waitlist decision making, as well as for identifying the highest performing students for recognitions, fellowships, scholarships and other meritorioius recognition. There are benefits also that may be derived from using situational judgment tests to predict student outcomes as measured by GPA and retention the following semester. Having this information enhances an institution's ability to target student support services to prevent student departures.

In conclusion, a shift in focus from a performance culture to one of mastery is needed in order to promote an enhanced, inclusive sophomore year ethos that is primarily focused on student engagement and student learning is an important direction for future explorations. There are significant opportunities to improve institutional supports that may be offered to improve and insulate the sophomore year experience. Using the regression models from this study, first year and admission variables as well as admission data may be added to the regression formula to predict the performance of the rising class of sophomore students during the next sequential semester. The application of these research findings may help promote a better institutional

response to the first year and provide additional supports to students transitioning from the freshman to the sophomore year. Situational Judgement Tests provide institutions with a unique opportunity to interpret existing archival data, while combining machine learning and data mining techniques to enhance insitutional responsiveness to the sophomore year population in particular.

SAT Math scores were negatively correlated with Spring retention. This finding suggests a need to re-evaluate the importance of standardized tests. It is recommended that institutions administer SJTs incrementally throughout the Undergraduate experience to help enhance institutional response while improving retention outcomes. This innovation provides a cost effective, systematic method that draws upon computers and technology and is attractive to institutions looking for viable and sustainable methods. Logistic regression models provide institutions with a formula to predict and respond according to the student needs providing an open window into students affective dimensions which are rarely gauged.

REFERENCES

Ackerman, P., Kanfer, R., & Beier, M. E. (2013). Trait complex, cognitive ability, and domain knowledge predictors of baccalaureate success, STEM persistence, and gender difference. *Journal of Educational Psychology, 105*(3), 911-927.

Ackerman, P., Chamorro-Premuzic, T., & Furnam, A. (2011). Trait complexes and academic achievement: Old and new ways of examining personality in educational contexts. *British Journal of Educational Psychology, 81*(Pt 1), 27-40. https://doi.org/10.1348/000709910X522564

April, K. A., Dharani, B., & Peters, K. (2012). Impact of locus of control expectancy on level of well-being. *Review of European Studies, 4*, 124-137. https://doi.org/10.5539/res.v4n2p124

Ashworth, V., Fung, K., & Shaw, R. (2014). National University of Singapore Situational Judgment Test and Focused Skills Assessment: Analysis & Evaluation. In Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. Medical Teacher, 38(1), 3–17. https://doi-org.libproxy.temple.edu/10.3109/0142159X.2015.1072619

Boivin, M., Fountain, G. A., & Baylis, B. (2000). Meeting the challenges of the sophomore year. In L. A. Schreiner & J. Pattengale. (Eds.), *Visible solutions for invisible students: Helping sophomores succeed (Monograph 31)* (pp. 118). University of South Carolina, National Resource Center for The First-Year Experience and Students in Transition.

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology, 81*(1), 78–96.

Cheng, D. X. (2004). Students' sense of campus community: What it means and what to do about it. *NASPA Journal, 41*, 216-232.

Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters, 115*(1), 11-15.

Culpepper, S. A. & Davenport, E. C. (2009). Assessing differential prediction of college grades by race/ethnicity with a multilevel model. *Journal of Educational Measurement, 46*(2), 220-242. https://doi.org/10.1111/j.17453984.2009.00079.

de Leng, W. E., Stegers-Jager, K. M., Born, M. P., & Themmen, A. P. (2018). Integrity situational judgement test for medical school selection: Judging 'what to do'versus 'what not to do.' *Medical Education, 52*(4), 427-437.

Delaney, L., Harmon, C., & Ryan, M. (2013). The role of noncognitive traits in undergraduate study behaviors. *Economics of Education Review, 32,* (2013), 181-195. https://doi.org/10.1016/jeconedurev.2012.07.009

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* Sage

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality & Social Psychology, 85*(3), 541–553. https://doi.org/10.1037/00223514.85.3.541

Graunke, S. S., & Woosley, S. A. (2005). An exploration of the factors that affect the academic success of college sophomores. *College Student Journal, 39*, 367-376.

Guenole, N., Chernyshenko, O. S., & Weekly, J. (2017). On Designing Construct Driven Situational Judgment Tests: Some Preliminary Recommendations. *International Journal of Testing, 17*(3), 234–252. https://doi.org/10.1080/15305058.2017.1297817

Gump, S. E. (2007). Classroom research in a general education course: Exploring implications through an investigation of the sophomore slump. *The Journal of General Education, 56*,105-125.

Hsieh, P., Sullivan, J. R., & Guerra, N. S. (2007). A closer look at college students: Self-efficacy and goal orientation. *Journal of Advanced Academics, 18*(3), 454–476. https://doi.org/10.4219/jaa-2007-500

Hotchkiss, J., Moore, R., & Pitts, M. M. (2006). Freshman learning communities, college performance, and retention. *Education Economics, 14*(2), 197–210. https://doi.org/10.1080/09645290600622947

Hunter, M. S., Tobolowsky, B. F., Gardner, J. N., Evenback, S. E., Pattengale, J. A., Schaller, M. A., & Schreiner, L. A. (2010). *Helping sophomores succeed: Understanding and improving the second-year experience. Jossey-Bass Inc.*

Kleiss, H. O., & Green, B. A. (2010). *Statistical concepts for the behavioral sciences* (4th ed.). Allyn and Bacon

Komarraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. *Learning & Individual Differences, 24*, 103–109. https://doi.org/10.1016/j.lindif.2012.12.007

Kreiter, C. D., & Axelson, R. D. (2013). A perspective on medical school admission research practice over the last 25 years. *Teaching and Learning in Medicine, 25*, 50-56.

Kuh, G. D., & Hu, S. (2001). The effects of student-faculty interaction in the 1990s. *Review of Higher Education, 2* , 309-332.

Lee, K., & Leonard, R. (2009). Destination unknown: Addressing the sophomore slump. *NACE Journal, 70*(2), 27–31.

Lievens, F., Buyse, T., & Sackett, P. R. (2005). The operational validity of a video-based situational judgment test for medical college admissions: Illustrating the importance of matching predictor and criterion construct domains. *Journal of Applied Psychology, 90*, 442-452.

McAbee, S. S., & Oswald, F. L. (2013). The criterion-related validity of personality measures for predicting GPA: A meta- analytic validity competition. *Psychological Assessment, 25*(2), 532-544. https://doi.org/10.1037/a0031748

McDaniel, M. A., & Whetzel, D. L. (2007). Situational judgment tests. *Applied Measurement: Industrial Psychology in human resources management*, 235-257.

McDaniel, M. A., & Whetzel, D. L. (2005). Situational judgment test research: Informing the debate on practical intelligence theory. *Intelligence, 33*(5), 515–525. https://doi.org/10.1016/j.intell.2005.02.001

Midgley, C., Kaplan, A., & Middleton, M. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? *Journal of Educational Psychology, 93*(1), 77–86. https://doi.org/10.1037/00220663.93.1.77

Milsom, C., & Yorke, M. (2015). Defining the second year. In C. Milsom, M. Stewart, M. Yorke, & E. Zaitseva (Eds.), *Stepping up to the second year at university: Academic, psychological and social dimensions* (pp. 14–23). Abingdon: Routledge.

Milsom, C., Stewart, M., Yorke, M., & Zaitseva, E. (2015). *Stepping up to the second year at university: Academic, psychological and social dimensions*. Abingdon: Routledge.

Mobarakeh, M. R. V., Juhari, R., Yaacob, S. N., & Redzuan, M. (2015). Locus of control and psychological well-being among Iranian adolescent migrants in Kuala-Lumpur, Malaysia. *American International Journal of Research in Humanities, Arts and Social Sciences, 10*, 310-313.

Myers, D. G., & Diener, E. (1995). Who is happy? *Psychological Science, 6*, 10-19. https://doi.org/10.1111/j.1467-9280.1995.tb00298.x

Nagda, B. A., Gregerman, S. R., Jonides, J., von Hippel, W., & Lerner, J. S. (1998). Undergraduate student-faculty research partnerships affect student retention. *Review of Higher Education, 22*(1), 55-72.

Nieves, W. (1991). *The relationship between minority student self-esteem and student retention at an urban community college* (Publication No. 991026642929704646) [Doctoral dissertation, Rutgers University]. ProQuest Dissertations & Theses Global.

O'Connell, M. S., Hartman, N. S., McDaniel, M. A., Grubb III, W. L., & Lawrence, A. (2007). Incremental validity of situational judgment tests for task and contextual job performance. *International Journal of Selection and Assessment, 15*(1), 19-29.

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure an situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89*, 187–207. https://doi.org/10.1037/0021-9010.89.2.187

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research.* Jossey-Bass.

Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education, 50*(1), 36–60. https://doi.org/10.1111/medu.12817

Patterson, F., Carr, V., Zibarras, L., Burr, B., Berkin, L., Plint, S., Irish, B., & Gregory, S. (2009). New machine-marked tests for selection into core medical training: Evidence from two validation studies. *Clinical Medicine 9*, 417-420.

Plominski, A. P., & Burns, L. R. (2018). An investigation of student psychological wellbeing: Honors versus nonhonors undergraduate education. *Journal of Advanced Academics, 29*(1), 5-28.

Richardson, M., Abraham, C. & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387. https://doi.org/10.1037/a0026838.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rodríguez-Muñiz, L. J., Bernardo, A. B., Esteban, M., & Díaz, I. (2019). Dropout and transfer paths: What are the risky profiles when analyzing university persistence with machine learning techniques? *PLoS ONE, 14*(6), 1–20. https://doi.org/10.1371/journal.pone.0218796

Sanchez-Leguelinel, C. (2008). Supporting "slumping" sophomores: Programmatic peer initiatives designed to enhance retention in the crucial second year of college. *College Student Journal, 42*, 637-646.

Sartania, N., McClure, J. D., Sweeting, H., & Browitt, A. (2014). Predictive power of UKCAT and other pre-admission measures for performance in a medical school in Glasgow: a cohort study. *BMC Medical Education, 14*(1), 1-10.

Schaller, M. A. (2005). Wandering and wondering: Traversing the uneven terrain of the second college year. *About Campus, 10*(3), 17-24. https://doi.org/10.1002/abc.131

Schreiner, L. A., & Pattengale, J. (Eds.). *Visible solutions for invisible students: Helping sophomores succeed. Monograph no. 31.* (2000). National Resource Center for the First-Year Experience and Students in Transition, University of South Carolina.

Schreiner, L. A., Louis, M. C., & Nelson, D. D. (2012). *Thriving in transitions : A research-based approach to college student success.* National Resource Center for The First Year Experience.

Senko, C., & Harackiewicz, J. M. (2002). Performance goals: The moderating roles of context and achievement orientation. *Journal of Experimental Social Psychology, 38*(6), 603-610.

Sharma, S., Gangopadhyay, M., Austin, E., & Mandal, M. K. (2013). Development and validation of a situational judgment test of emotional intelligence. *International Journal of Selection & Assessment, 21*(1), 57–73. https://doi.org/10.1111/ijsa.12017

Stewart, M., & Darwent, S. (2014). *Psychological orientations to learning in the second year. Stepping up to the second year at university* (pp. 59-73). Routledge.

Tinto, V. (1975). Drop out from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*, 95-112.

Tinto, V. (1982). Limits of theory and practice in student attrition. *Journal of Higher Education, 53,* 687-700.

Tobolowsky, B. F., & Cox, B. E. (2012). Rationalizing neglect: An institutional response to transfer students. *Journal of Higher Education, 83*(3), 389–410. https://doi.org/10.1080/00221546.2012.11777249

Tower, M., Blacklock, E., Watson, B., Heffernan, C., & Tronoff, G. (2015). Using social media as a strategy to address "sophomore slump" in second year nursing students: A qualitative study. *Nurse Education Today, 35*, 1130-1134. https://doi.org/10.1016/j.nedt.2015.06.011

Wang, X., & Kennedy-Phillips, L. (2013). Focusing on the sophomores: Characteristics associated with the academic and social involvement of second-year college students. *Journal of College Student Development, 54*, 541-548. https://doi.org/10.1353/csd.2013.0072

Webb, O. J., & Cotton, D. R. E. (2019). Deciphering the sophomore slump: Changes to student perceptions during the undergraduate journey. *Higher Education, 77*(1), 173–190.

Westring, A. F., Oswald, F., Schmitt, N., Drzakowski, S., Imus, A., Kim, B., & Shivpuri, S. (2009). Estimating trait and situational variance in a situational judgment test. *Human Performance, 22*(1), 44–63. https://doi.org/10.1080/08959280802540999

Whetzel, D .L, & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, 19*(3) 188-202. https://doi.org/10.1016/j.hrmr.2009.03.007

Wilson, K. J., & Crowe, M. (2010). Undergraduate research: A powerful pedagogy to engage sophomores. In M. Hunter, B. Tobolowsky, J. Gardner, S. Evenbeck, J. Pattengale, M. Schaller, & L. A. Schreiner (Eds.), *Helping sophomores succeed: Understanding and improving the second-year experience* (pp. 177-188). Jossey-Bass.

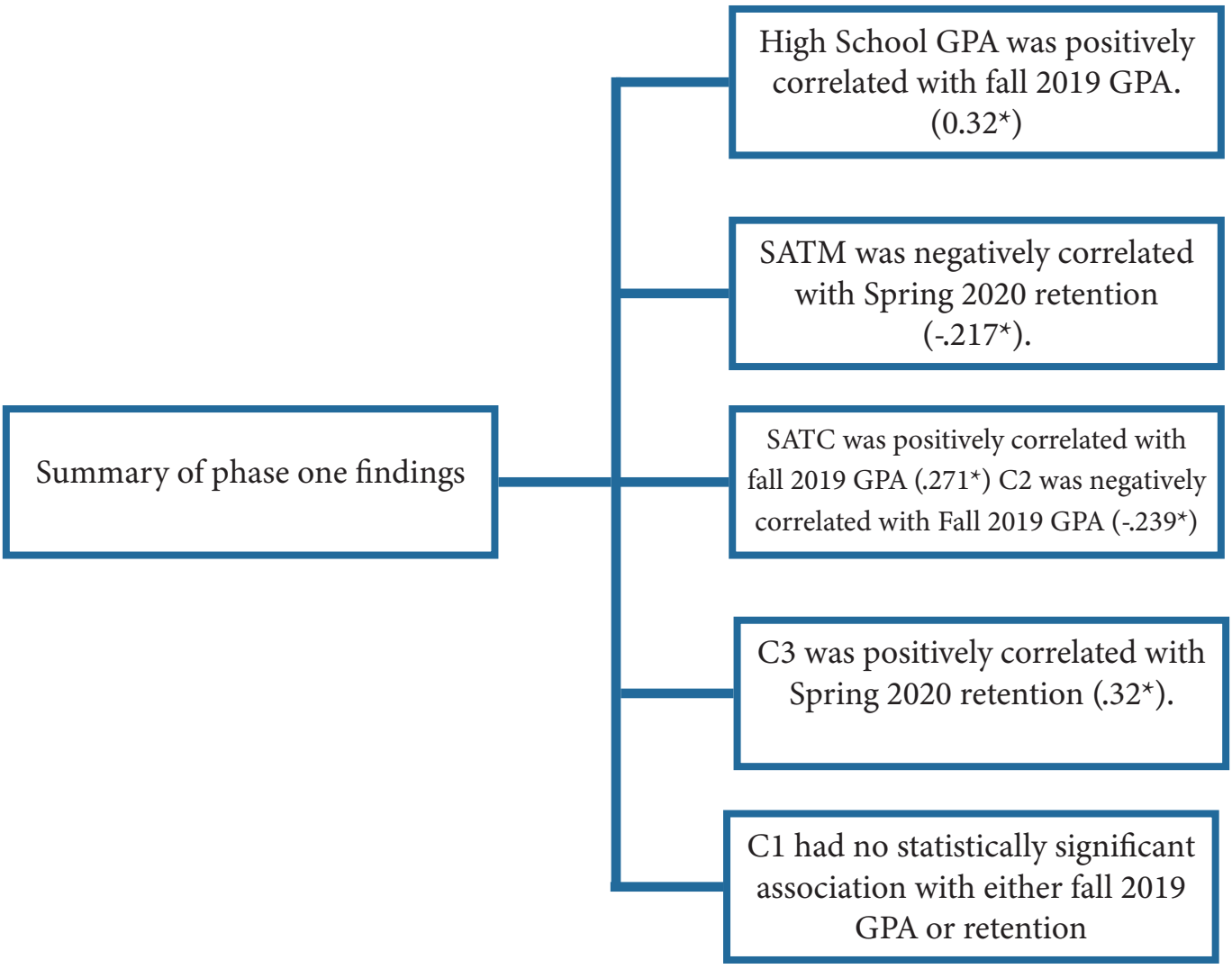
Yorke, M., Hoekstra, P., & Turnbull, W. (2015). Maximizing the use of institutional data, 84-103, In Milsom, C., Stewart, M., Yorke, M., & Zaitseva, E. (2015). *Stepping up to the second year at university: Academic, psychological, and social dimension.. Taylor & Francis: New York*, 1-131.

Zaitseva, E., Milsom, C., & Stewart, M. (2013). Connecting the dots: Using concept maps for interpreting student satisfaction. *Quality in Higher Education, 19*(2), 225–247.

APPENDICES

Appendix A: Figure 1: Summary of Results
Appendix B: Table 1: Number of Represented in the Study by Department
Appendix C: Table 2: Summary Statistics for Selected Study Variables
Appendix D: Table 3: The Frequency Counts for Retention Spring 2020
Appendix E: Table 4: Correlation Matrix for Fall 2019 GPA & Survey Factors
Appendix F: Table 5: Point-Biserial Correlations Between Survey Factors & Fall 2019 Retention
Appendix G: Table 6: Multiple Regression Model with Fall 2019 GPA as Dependent Variable
Appendix H: Table 7: The Logistic Regression Model for Spring 2020 Retention
Appendix I: Table 8: Intercorrelations Between Survey Factors and Fall 2019 GPA

Appendix A
Summary of Results



Appendix B
Number of Students Represented in the Study by Department

Table 1

Number of Students Represented in the Study by Department

DEPARTMENT	*n*	PERCENT
BIOLOGY	31	35%
EDUCATION	34	32%
SOCIOLOGY/SW/CRIMINAL JUSTICE	21	21%
PSYCHOLOGY	10	10%
UNSPECIFIED	2	2%
TOTAL	n=98	100%

Note. n=98

Appendix C
Summary Statistics for Selected Study Variables

Table 2

Summary Statistics for Selected Study Variables

VARIABLE	*n*	*M*	*SD*	95% CI
FALL 2019 GPA	79	3.12	0.58	[2.98, 3.25]
C1-MASTERY	79	21.96	3.73	[21.13, 22.80]
C2-APPROACH	79	27.63	4.12	[26.71, 28.56]
C3-AVOID	79	17.11	5.26	[15.94, 18.29]
HSGPA	79	3.31	0.41	[3.23, 3.40]
NSATERW	79	523.67	73.91	[507.12, 540.22]
NSATM	79	499.49	72.82	[483.18, 515.80]

Note. n = Sample size; M = mean; SD = Standard Deviation; CI = Confidence Interval for the mean. C1-Mastery = Developing and improving abilities; C2-Approach = demonstrating ability; C3Avoid = hiding lack of ability. HSGPA = high school GPA ; NSATERW = New SAT; NSATM = SAT Math.

Appendix D
The Frequency Counts for Retention Spring 2020

Table 3

The Frequency Counts for RetSp2020

VARIABLE	FREQUENCY	RELATIVE FREQUENCY (%)
RetSp2020		
YES	88	89.8
NO	10	10.2

Note. Relative Frequency = percent of total.

Appendix E
Correlation Matrix for Fall 2019 GPA & Survey Factors

Table 4

Correlation Matrix for Fall 2019 GPA and Survey Factors

		Grade fall 2019	C1-Mastery	C2-Approach	C3-Avoid
Grade fall 2019	Pearson Correlation	1	-.163	-.239*	.052
	Sig. (2-tailed)		.126	.024	.632
	N	90 -	89	89	89 -
C1-Mastery	Pearson Correlation	.163	1	.607**	.547**
	Sig. (2-tailed)	.126		.000	.000
	N	89	97	97	97
C2-Approach	Pearson Correlation	-.239*	.607**	1	-.524**
	Sig. (2-tailed)	.024	.000		.000
	N	89	97	97	97
C3-Avoid	Pearson Correlation	.052	-.547**	-.524**	1
	Sig. (2-tailed)	.632	.000	.000	
	N	89	97	97	97

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Note. C1-Mastery = ; C2-Approach = ; C3-Avoid = .
Numbers in parentheses represent correlation sample sizes.
*p < .05. **p < .01.

Appendix F
Point-Biserial Correlations Between Survey Factors & Fall 2019 Retention

Table 5

Point-Biserial Correlations Between Survey Factors and Fall 2019 Retention

	FALL 2019 RETENTION	C1-MASTERY	C2-APPROACH	C3-AVOID
Fall 2019	--	.07	.04	.19
Retention		(97)	(97)	(97)

Note. C1-Mastery = Developing and improving abilities; C2-Approach = demonstrating ability; C3-Avoid = hiding lack of ability. Numbers in parentheses represent correlation sample sizes.

Appendix G
Multiple Regression Model with Fall 2019 GPA as Dependent Variable

Table 6

Multiple Regression Model with Fall 2019 GPA as Dependent Variable

PREDICTOR	B	SE B	b
INTERCEPT	1.89*	0.73	0
C2-APPROACH	-0.03*	0.02	-.22
HSGPA	0.41*	0.16	.28
NSATERW	0.001	0.001	.17

Note. n = 77. B = unstandardized beta coefficients.
SE = standard error of beta.
b = standardized beta coefficients.
$F(3,73) = 5.89$, $p < .001$, $R2 = .20$.
*$p < .05$.

Appendix H
The Logistic Regression Model for Spring 2020 Retention

Table 7

The Logistic Regression Model for RetSp2020

VARIABLE	*B*	*SE*	*Exp(B)*	95% CI for *Exp(B)*
CONSTANT	-11.01	6.57		
C1-MASTERY	0.16	0.16	1.17	(0.86, 1.61)
C2-APPROACH	0.12	0.16	1.13	(0.83, 1.55)
C3-AVOID	0.32*	0.15	1.38	(1.03, 1.84)
HSGPA	2.10	1.19	8.16	(0.79, 84.70)
NSATERW	0.02*	0.01	1.02	(1.00, 1.05)
NSATM	-0.03*	0.01	0.97	(0.94, 0.99)

Note. n = 84. R2 = .50 (Nagelkerke). Model c2(6) = 24.93, p < .001.
*p < .05.

Appendix I
Intercorrelations Between Survey Factors and Fall 2019 GPA

Table 8

Intercorrelations Between Survey Factors and Fall 2019 GPA

AVOID	FALL 2019 GPA	C1-MASTERY	C2-APPROACH	C2-APPROACH
FALL 2019 GPA	--	-.16 (89)	-.24* (89)	.05 (89)
C1-MASTERY		--	.61** (97)	-.55** (97)
C2-APPROACH			--	-.52** (97)
C3–AVOID				--

Note. C1-Mastery = Developing and improving abilities; C2-Approach = demonstrating ability; C3-Avoid = hiding lack of ability.

Numbers in parentheses represent correlation sample sizes.
*p < .05. **p < .01.

ABOUT THE AUTHOR

Amml Hussein, Ed.D., is an assistant professor in Temple University's School of Social Work. A seasoned social worker, educator and researcher, Dr. Hussein brings more than 12 years of service experience in several practice domains including medical social work, utilization management, complex managed care, psychiatric crisis stabilization, curriculum and instruction, and instructional design.

Dr. Hussein is committed to improving the college student experience, with particular emphasis on the second and third years, and increasing institutional responsiveness to students. Her teaching philosophy is centered in activated learning, where she extends course concepts beyond the classroom and situates dialogues in broader international and global contexts. Some of Dr. Hussein's research interests include integrating anti-oppressive, empowerment models, and trauma informed care into curriculum, social and emotional predictive models and the science of student success, health promotion strategies geared to alleviating health disparities, and enhancing adult learner engagement infusing innovative teaching modalities and technology to enhance knowledge acquisition.

amml@temple.edu