Journal of Education and Social Development, 2020, 4–1
July. 2020, pages 1-9
doi: 10.5281/zenodo.3911972
http://www.ibii-us.org/Journals/JESD/
ISBN 2572-9829 (Online), 2572-9810 (Print)

Connecting STEAM Classroom Observation Data to Student Achievement Data: An Empirical Perspective

Carla Thompson, Ed. D.[1], Karen Barber, Ed. D.[2], Joan Nkansah, Ed.S.[1], Bright Da-Costa Aboagye, Ed.S.[1], Sandra Ayivor, Ed.S.[1], Kym Atwood, Ed.S.[1], and Candy Hurst, B.S.[1]

[1]Department of Educational Research and Administration, University of West Florida, Florida, United States of America

[2]Department of Federal Programs, Santa Rosa County Public School District, Santa Rosa County, Florida, United States of America

*Email: cthompson1@uwf.edu

Received on 04/22/2020; revised on 06/03/2020; published on 06/28/2020

Abstract

Emerging STEAM (Science, Technology, Engineering, Art, and Mathematics) classrooms in K-12 schools across the United States have prompted researchers to begin questioning the impact of STEAM instruction on student achievement. This study examined the relationship between STEAM classroom instructional environment measures using a quantitative STEAM classroom observation form developed and piloted by the researchers and students' achievement scores in mathematics and language arts using standardized achievement tests. Researchers captured multifaceted STEAM instructional strategies associated with creativity, critical thinking, communication, and collaboration skills using a numeric observation scale associated with four levels of intensity recorded by university trained research assistants. A series of Spearman Rho correlation procedures were performed to examine the relationship between the classroom observation data (independent variables) and state standardized achievement scores (dependent variables). These data were empirically connected to students' achievement scores derived from state standardized testing data. Results provide an empirically-based rationale for assessing STEAM instructional practices using classroom observational data. Implications of empirically connecting STEAM instructional practices to student achievement outcomes may offer an action plan and methodology for school districts to use for capturing empirical evidence, thereby assisting schools and school districts grappling with STEAM funding challenges. Implications of the study may also provide evidence for researchers for determining: (a) effective STEAM instructional strategies; (b) purposeful STEAM professional development topics; and (c) strategic workforce preparation skills necessary for moving STEAM forward into a global community of learners.

Keywords: STEAM Education, Classroom Observation, Standardized Testing, Spearman Rho Correlation

1 Introduction and Background

The United States educational focus of implementing STEM (Science, Technology, Engineering, and Mathematics) instructional programs in K-12 schools has been determined by the President and congress to provide high quality educational STEM programs for the purpose of identifying the United States as the "global leader in STEM literacy, innovation, and employment" with "all Americans having lifelong access to high quality STEM education" (Office of the President of the United States, 2018, p. 4). Moving this commitment forward to include STEAM (adding the Arts) is also supported by the national movement for STEM and is specifically focused on developing students' problem-solving abilities, writing, communication, language arts skills, mathematics, engineering, and science capabilities. These efforts may move learners toward a global STEAM community. Recent research literature findings focused on the implementation of STEAM into K-12 schools within the United States have earmarked several areas of needed research activities. Examples of research needs are highlighted within the extensive literature review of STEAM education conducted by Perignat and Katz-Buoincontro (2019) who provided a dual perspective of the purpose and need for STEM/STEAM education aligned with the current study purpose: (a) to enhance the "importance of advancing learning in STEM disciplines and engaging minority and female students in STEM subjects, increasing interest in STEM fields and developing skills necessary for STEM careers." (p. 34) and (b) to integrate "domain general-skills such as perspective-taking creative and problem-solving skills, knowledge transfer across disciplines, and/or encouraging students to explore and experience new ways of knowing" (p. 34). The added content of the Arts or the "A" within the context of the STEAM literature review conducted by Perignat and Katz-Buoincontro (2019), examined by multiple studies determined the "A" to be the impetus for moving STEM learning toward creative problem-solving and creative thinking, for example, connecting to the arts is "characterized as one of the benefits or learning outcomes of STEAM education." (p. 38). Additionally, the extensive literature review conducted by Perignat and Katz-Buoincontro (2019) reported a need not yet addressed by researchers (i.e., "scholars posit that STEAM is a model for enhancing creativity, critical thinking, and other skills, however, there is little evidence to support this notion" [p. 40]). Literature identifying the need for examining the skills represented/observed within STEAM classrooms and the desire to determine the degree of effectiveness of the integration of science, technology, engineering, arts, and mathematics relative to student outcomes within schools provided the overriding impetus for initiating the current study. The current study empirically explored connections between observed STEAM classroom instructional approaches involving creativity, critical thinking, and other key skills relative to student performance outcomes using structured classroom observations as the key element for measuring instructional environments and students' standardized tests scores as measures of student performance outcomes. Prior research efforts aimed at examining the content and instructional approaches included within STEAM classrooms coupled with previous research projects aimed at exploring methods for determining the effectiveness of STEAM approaches have provided insight for the current study focused approach. A summary of selected prior research approaches for consideration relative to the current study focus is presented in the following chronological listing: Miller and Knezek (2013) advocated for the inclusion of collaboration, creativity, and problem solving as critical topics for classroom instruction strategies; Bush, Cook, Ronau, Rakes, Mohr-Schroeder, and Saderholm (2016) identified several characteristics for effectively assessing the implementation of STEM education in terms of whole group engagement, including "analysis of classroom practices, analysis of student work and learning, and on-going engagement over time" (p. 109); Quigley, Herro, and Jamil (2017) provided a conceptual model of STEAM (Science, Technology, Engineering, Art, and Mathematics) relative to classroom teaching activities and emphasized the need for research evidence to explain "how instructional approaches must shift to enact effective STEAM teaching" (p. 2); Harper (2017), representing the Association of Supervision and Curriculum Development (ASCD) posited a "Call to Action" for teachers and leaders to employ "active inquiry, critical thinking, and problem-solving skills" as key tenets of STEAM teaching and learning environments (p. 70). This focused perspective has prompted educational researchers to seek empirical methods for determining effective STEAM instructional strategies. Research approaches posited by Perignat and Katz-Buoincontro (2019) included the need for research focused on understanding "the implications of STEAM in practice" (p. 35) and the need for researchers to examine thinking skills reflective of STEAM learning, i.e., critical thinking, creativity, problem-solving, and other types of thinking skills. Additional research findings focused on the need for researchers to examine two perspectives for determining outcomes from STEAM education: (1) the integrated disciplines of science, technology, engineering, art, and mathematics (STEAM) and the pedagogies associated with the integration of these disciplines and (2) the methods for assessing student learning outcomes from the integration of these various disciplines. Nadelson and Seifert (2017) defined integrated STEM as "the seamless amalgamation of content and concepts from multiple STEM disciplines" (p. 221). These researchers also indicated the need for a shift toward a problem-based curriculum "focused on student-centered learning and less on knowledge conveyance" (p. 223). Herro and Quigley (2017) advocated for STEAM instruction to be focused on Problem-Based Learning (PBL) and emphasized the need for teachers to become adequately prepared in PBL through professional development programs. Constantino (2018) supported integrating creative inquiry as an instructional method through art and design within STEAM classrooms for promoting a transdisciplinary curriculum model. Harper (2017) joined multiple authors, researchers, and educators in supporting active teaching and learning derived from a cognitive, experiential foundation involving students as the key elements in the learning process. Hackathorn, Solomon, Blankmeyer, Tennial, and Garczynski (2011) found active learning activities resulted in high performance outcomes for college students while lecture-driven instruction produced low performance outcomes for college students. Focused research studies identifying various approaches for assessing STEAM learning outcomes have provided descriptive information pertinent to the current study procedures. Hora and Ferrare (2013) presented information concerning the use of classroom observation techniques in postsecondary settings as applications for assessing STEAM classroom instructional approaches. Stuhlman, Hamre, Downer, and Pianta (2018) with the Center for Advanced Study of Teaching and Learning (CASTL) at the University of Virginia provided a strong rationale for the use of standardized, reliable, and validated observational tools substantiated by the following statement: "Using reliable and valid tools to observe and rate classrooms provides a research-based mechanism for achieving a second step toward systematic increases in educational quality" (p. 4). Duke and Halvorsen (2017) examined the impact of using PBL instructional methods on elementary students in social studies and utilized three assessments of achievement (social studies content, informational writing, and informational reading) aligned with state standards as the achievement measures of student outcomes. Emerging STEAM (Science, Technology, Engineering, Art, and Mathematics) programs and classrooms in K-12 schools have prompted researchers to begin questioning the impact of STEAM instruction on student achievement. Limited research findings exist to explain how the arts, and mathematics, including Emerging STEAM (Science, Technology, Engineering, Art, and Mathematics) programs and classrooms in K-12 schools have prompted researchers to begin questioning the impact of STEAM instruction on student achievement. Limited research findings exist to explain how STEAM classroom instruction impacts student achievement in the sciences, technology, engineering language arts and other common test topics included in standardized testing across K-12 schools within the United States.

2 Methods

The current study explored connecting classroom observations of instructional strategies observed within STEAM classrooms in Grades 3 to 10 to achievement scores of students within the observed classrooms. The current study examined STEAM classroom instructional strategies within K-12 schools using trained researchers as classroom STEAM observers and paired these classroom instructional observations with students' scores on standardized achievement tests within a specific school district in the southeast region of the United States. The major research question guiding the current quantitative study was the following: RQ - How is the implementation of STEAM education within K-12 schools as measured by quantitative classroom observations using trained university observers related to students' achievement scores for students in Grades 3 to 10 using state mandated standardized achievement testing measures? The independent variable (STEAM education implementation) was measured by the STEAM classroom observation form data from more than 100 observations collected in classrooms by trained university observers and the dependent variable was the resulting scores of classrooms of students on standardized tests. The unit of measure used for the current study was the STEAM classroom. Instrumentation used for measuring the independent variable (Classroom Observation Form data) and dependent variable (state standardized test scores) with the respective specific data types aligned with the major research question posited for the current study are depicted in Figure 1 and Table 1.

Classroom Observation Form Focus Areas	Classroom Observers[1]
1) Creative Preparation (CP)	1 =Descriptive
2) Critical Inquiry (CI)	2 =Emerging
3) Critical Thinking Integration (CTI)	3 =Developing
4) Critical Thinking: Problem Solving (CTP)	4 =Accomplished
5) Critical Thinking: Logical Thinking (CTL	
6) Communication: Data & Information Collection (CDI)	
7) Communication Argumentation (CA).	
8) Collaboration Team Work (CTW)	
9) Collaboration Investigation Skills (CIS)	

Figure 1. Classroom Observation Categories

Table 1. Exemplars, Sample Descriptors, and Partial Form Used for Determining Observation Rating Categories from the Classroom Observation Form Used in the Study

Classroom Observation Focus Area	Descriptive Rating = 1	Emerging Rating = 2	Developing Rating = 3	Accomplished Rating = 4
Creative Preparation	Lesson incorporates opportunities for students to investigate local and global issues, universal problems, and transdisciplinary ideas	Lesson guides student activity to support core ideas, practices, and academic content standards. Students follow directives to generate a solution.	The teacher designs classroom activities that link academic standards with transdisciplinary lessons that may involve local and global issues. Teacher asks students to generate new investigations.	Teacher designs authentic experiences of transdisciplinary student inquiry and design. Teacher encourages students to think outside the box. Students demonstrate progress in unique and creative ways through multiple mediums.
Creative Inquiry	Students are directly taught and expected to ask questions, identify problems, seek resources, and persevere in problem solving	Teacher directs inquiry using a limited set process. Students must use a question to begin an inquiry.	Teacher predominately initiates inquiry and guides students through differing models of inquiry. Students engage in their own inquiries.	Inquiry may be student directed. Teacher serves as a facilitator of inquiry. Students connect that many careers use varied types of inquiry models.
Critical Thinking Integration	Learning experiences are transdisciplinary in nature and focus on authentic content connections, and	Teacher plans multidisciplinary experiences focused on common themes but stays within the content using prompted	Teacher plans multidisciplinary learning experiences that link multiple content areas The teacher relates the	Teacher plans meaningful authentic opportunities for students to analyze real world relationships across

	current real-world problems	discussions using skills and content across two subject areas.	interdisciplinary nature of problems to real world situations.	content areas. Students explore trans-disciplinary connections. Teacher supports students in understanding real world problems.
Critical Thinking Problem Solving	Students are taught and expected to construct explanation, design solutions, and solve problems using textual and empirical evidence.	Teacher leads instruction on constructing explanations, designing solutions, and solving problems using evidence. Teacher guides students toward evidence.	Teacher guides students through activities for providing solutions based on evidence. Teacher directs conversations with students to determine legitimate solutions.	Teacher designs experiments and students provide multiple explanation and solutions based on evidence from a variety of sources. Students are expected to determine if there is enough evidence for legitimate conclusions.

The outcome measures used to determine the effectiveness of the implementation of STEAM education in grades 3 to 10 within a school district in the southeast region of the United States included scores on an annual state mandated test depicted in Figure 2.

STATE MANDATED TEST DATA CATEGORIES and Grade Levels

1) Integration of Knowledge/Ideas…… Elem
2) Key Ideas and Details…………… .. Elem
3) Language Editing Tasks…………… Elem
4) Craft and Structure………………… Elem
5) Text-based Writing …………… … Elem/Mid
6) Measurement/Data/Geometry……… Elem/Mid
7) Numbers/Operations/Fractions……… Elem/Mid
8) Operations/Algebraic Thinking………Middle
9) Ratio/Proportional Relations………… Middle
10) Operations/Algebraic Thinking… Middle
11) Expressions & Equations………… Middle
12) Geometry………………………… Middle
13) Operations/Algebra Thinking …… Middle
14) Statistics and Probability…………… High
15) The Number System………………… High
16) Functions…………………………… High
17) Statistics/Probability/Number Systems… High
18) Algebra Modeling……………………… High
19) Functions/Modeling…………………… High
20) Statistics/Number Systems …………………High
21) Congruence/Triangles/Trigonometry………High
22) Circles/Geometric Measures/Equations… Mid/High
23) Modeling with Geometry……………………High

Figure 2. State Mandated Test Categories & Grade Levels for Each Test

A school district within the southeast region of the United States requested research/evaluation services and contracted with the neighboring university research and evaluation team to conduct an extensive research and evaluation of the implementation of STEAM education within the district. The county containing the school district used in the study represents a semi-rural population of approximately 175,000 people with a low-income farming community in the northern part of the county and a gulf-shore, white sands, ocean vacation high-income resort community in the southern part of the county. The school district/county extends from the top of the state with the state line as the top border of the county to the Gulf of Mexico as the southern border of the county also aligning with the school district borderlines. The school district identified the specific schools with STEAM classrooms/laboratories for participation in the current study. The study was comprised of 26 public schools including six high schools, seven middle schools, and 13 elementary schools. STEAM Classrooms representing the 26 schools focused within the study included both STEAM designated classrooms or STEAM laboratories located within each school. Six observers trained to use the STEAM Classroom Observation Form were comprised of graduate students enrolled in the University of West Florida Doctor of Education (Ed. D.) Curriculum and Instruction program of study. Observers were each randomly assigned to four or five of the 26 schools within the school district and were required to complete two fifteen-minute observations per week in each of their schools for 10 weeks in the fall semester and 10 weeks in the spring Semester. Two of the six observers were assigned to five schools with four observers each assigned to six schools respective of the locations of the schools to allow for driving and time considerations. Interrater reliability (Halgren, 2012) was performed among the six observers relative to the overall use of the STEAM Classroom Observation Form using four weeks of classroom observations with a resulting alignment coefficient of .87 ($p = .02$) using an inter-class correlation (ICC) coefficient procedure appropriate for use when utilizing ordinal ratings and more than two observers (Cicchetti, 1994). Research associates and research assistants engaged in the project assisted with the data entry for N= 914 total observations over a 20-week time period and aligned observation data with the appropriate schools and STEAM classrooms. Ordinal data from the 914 classroom observations were then examined by researchers entering data and the largest occurring data level (1, 2, 3, or 4) was selected for the independent variable to allow for a one-to-one correspondence for depicting a rating (descriptive, emerging, developing, and accomplished) for each of the school's individual STEAM classrooms relative to the eight observed classroom characteristics measured (creative inquiry, critical thinking integration, critical thinking problem solving, critical thinking logical thinking, communication data and information collection, communication argumentation, collaboration team work, collaboration investigation skills, and creative preparation). The approval for accessing student end-of-year state standardized test scores was provided by the school district with appropriate provisions for maintaining anonymity of students' names and/or identifiers. Student data were paired with STEAM classroom observation data based on school names and grade levels and respective of STEAM classrooms within the schools. Aligning end-of-year standardized test scores by school with STEAM classroom observation data by school created a paired database for empirically examining the impact of STEAM instructional strategies (as measured by STEAM classroom observation data) on student performance levels (as measured by students' state-mandated standardized test scores). The unit of analysis used for the analysis of data procedures for the study was the individual STEAM classrooms aligned with each of the 26 schools participating in the study. The overriding research question guiding the study was the following: How is the implementation of STEAM education within K-12 schools as measured by quantitative classroom observations using trained university observers related to students' achievement scores for students in Grades 3 to 10 using state mandated standardized achievement testing measures. The focused form of the research question applicable for quantitative analyses is the following question: Is there a statistically significant relationship between STEAM classrooms' ratings (Descriptive, Emerging, Developing, and Accomplished) relative to nine observation categories (see Figure 1) and aligned state test scores of students within STEAM classrooms. The quantitative data analysis procedure appropriate for examining relationships between STEAM classroom observation data and students' end-of-year standardized test score data was the Spearman Rank Order correlation analysis procedure with classroom observation data as the independent variable (ordinal level data) and standardized test scores as the dependent variable (interval level data). Prior to performing the Spearman correlational analyses, the assumptions associated with the Spearman correlational analyses were examined for alignment. The assumptions include the following: (a) data must be at least ordinal level data and (b) the scores on one variable must be monotonically related to the other variable (Warner, 2013). The independent variable is ordinal level data and the dependent variable is interval level data with a one-to-one correspondence or monotonic data entry from each of the two variables satisfying the assumptions of Spearman Rho Correlation. The Spearman Rho Correlation procedure was implemented using appropriate statistical software for the purpose of empirically determining a response to the major research question posited for the study, i.e., What is the relationship between the implementation of STEAM education within K-12 schools as measured by quantitative classroom observations using trained university observers relative to students' achievement scores using state mandated standardized testing measures? Researchers aligned the ordinal data depicted by the independent variable (classroom observation data) with the interval level data depicted by the dependent variable (standardized test scores represented by classroom averages) for achieving the necessary data alignment measures for computing the Spearman Rho Correlation coefficient, thereby meeting the assumptions of the Spearman correlation procedure, i.e., each of the independent and dependent variables must be at least ordinal level data and the scores on one variable must be monotonically related to the other variable. The individual STEAM classrooms were identified as the unit of measure for the study, thereby identifying the independent variable as the classroom observation measures depicted as ordinal level data. The dependent variable, average scores by classroom on state achievement tests were depicted as interval level data. These two types of data were aligned with the monotonic relationship identified with the Spearman Rho Correlation procedure. A series of Spearman Rho Correlation procedures were performed using classroom observation data as the independent variables and state standardized achievement scores as the dependent variables with the unit of measure for the analyses as the individual STEAM classrooms. Students in Grades 3-10 were involved in the study.

3 Results.

Resulting Spearman Rho correlation coefficients are reported in Tables 2 and 3 (see Appendix). Table 2 provides the resulting Spearman Rho values and associated significance levels for each of the classroom observation categories aligned with each of the standardized mathematics tests administered in grades 3 to 10 within the school district and Table 3 provides the analogous information (Spearman Rho correlation coefficients and significance values) for each of the classroom observation categories aligned with the various language arts standardized tests administered in grades 3 to 10 within the school district. The unit of measure was the classroom and the number of classrooms analyzed per standardized test differed somewhat by grade level and testing area. For example, Algebra and higher-level mathematics classes represented secondary grades (N= 6 classrooms) while language arts classes encompassed all of the classrooms (N= 20 classrooms) from grades 3 to 10.

3.2 Interpretation of Results

Examining interrelationships between key instructional elements or components involved in the teaching and learning of STEAM findings are depicted in the set of findings listed below with a substantial number of significant ($p < .05$ and $p < .10$) findings indicated in the content areas of mathematics and one significant ($p < .10$) finding within the area of language arts. A summary interpretation of these significant findings is provided in list form.

- *Creative Preparation (CP)*: No significant relationships were found between classroom observations of instruction documented as Creative Preparation and mathematics achievement scores or language arts achievement scores. These results may be interpreted in several ways: (a) observation data depicting examples of creative preparation may not align with key elements of achievement data within mathematics or language arts testing areas; (b) teachers focused on encouraging students to think outside the box may not be emphasizing rigid factual information within instruction; and (c) standardized testing items may not include items focused on allowing students to display creative preparation skills.
- *Creative Inquiry (CI)*: Creative Inquiry was found to be significantly related (Rho =.53, $p < .05$) to Number Operations and Base 10 standardized test scores. This result provides some evidence in support of Creative Inquiry as an effective instructional approach for enhancing arithmetic operations and Base 10 concepts. No significant relationships were found between any of the CI observation data and language arts achievement test scores.

- *Critical Thinking: Integration (CTI):* Critical Thinking Integration was found to be significantly related (Rho = .42, $p < .10$) to Number Operations and Base 10 standardized test scores lending some evidence in support of instructional practices encouraging critical thinking with integration as a positive influence in promoting number operations and Base 10 computation skills. No significant relationships were found between any of the CTI observation data and language arts achievement test scores.
- *Critical Thinking: Problem Solving (CTP):* Critical Thinking Problem Solving was found to be significantly (Rho = .42, $p < .10$) related to Number Operations and Base 10 standardized test scores lending some evidence in support of instructional practices encouraging critical thinking focused on problem solving as a positive influence in promoting number operations and Base 10 computation skills. No significant relationships were found between any of the CTP observation data and language arts achievement test scores.
- *Critical Thinking: Logical Thinking (CTL):* Critical Thinking Logical Thinking was found to be significantly related (Rho = .42, $p < .10$) to Number Operations and Base 10 standardized test scores. CTL was also found to be significantly related (Rho = .79, $p < .05$) to Algebra Functions, Algebra Modeling (Rho = .75, $p <.05$), and Algebra/Statistics/Number Systems (Rho = .75, $P<.05$). These results provide some empirical evidence for connecting instructional practices emphasizing critical thinking and logical thinking skills to positive student performance in number operations, algebra, and functions, and other advanced mathematical skill sets. No significant relationships were found between CTL and any of the language arts achievement test scores.
- *Communication: Data and Information (CDI):* No significant relationships were found between classroom observations of instruction documented as Communication with Data and Information and mathematics achievement scores or language arts achievement scores. These results may be interpreted in several ways: (a) observation data depicting communication data and information may not align with key elements of achievement data within mathematics and language arts testing areas; (b) teachers focused on encouraging students to think outside the box may not be emphasizing rigid factual information within instruction; and (c) standardized testing items may not include items focused on allowing students to demonstrate communicating with data and information skills.
- *Communication: Argumentation (CA):* Communication using argumentation was found to be significantly related (Rho = .48, $p<.05$) to mathematics achievement (Number Operations and Base 10). This result provides some evidence in support of communicating using argumentation as an effective instructional approach for enhancing student achievement in arithmetic operations and Base10 concepts.
- *Collaboration: Team Work (CTW):* No significant relationships were found between classroom observations of instruction documented as Collaborative Team Work and mathematics achievement scores or language arts achievement scores. These results may be interpreted in several ways: (a) observation data depicting examples of team work may not align with key elements of achievement data within mathematics or language arts testing areas; (b) teachers focused on encouraging students to think outside the box may not be emphasizing team work activities within instruction; and (c) standardized testing items may not include items focused on allowing students to exhibit team work skills.
- *Collaboration: Investigation Skills (CIS)*: Collaborative Investigation Skills were found to be significantly related (Rho = .37, $p < .10$) to text-based writing skills within the standardized language arts testing scores. This result may be evidence in support of using textbook and library resources for students within collaborative investigation activities. No significant relationships were found between CIS and mathematics achievement scores.

Conclusions

The study responded to the major research question: RQ - How is the implementation of STEAM education within K-12 schools as measured by quantitative classroom observations using trained university observers related to students' achievement scores for students in grades three to ten using state mandated standardized achievement testing measures? The independent variable (STEAM education implementation) was measured by the STEAM classroom observation form data from more than 100 observations collected in 20 elementary schools and six middle and high school schools by trained university observers and the dependent variable was the resulting scores of classrooms of students on standardized tests. The unit of measure used for the current study was the STEAM classroom. Instrumentation used for measuring the independent variable (Classroom Observation Form data) and dependent variable (state standardized test scores) produced nine significant relationships using the Spearman Rho Correlation Coefficient procedure. The following observed classroom instructional strategies: Creative Inquiry (CI), Critical Thinking Inquiry (CTI), Critical Thinking Problem-Solving (CPS), Critical Thinking Logical Thinking (CTL), Communication Argumentation (CA), and Collaboration Investigation Skills (CIS) were found to have significant relationships with classroom achievement scores in mathematics, algebra, advanced algebra skills, and text-based writing skills in language arts standardized testing results. Study results lend empirical evidence to the use of connecting classroom observation data to standardized testing data for empirically examining STEAM instructional strategies and effectiveness as represented by classroom observations of instruction and student standardized testing/achievement outcomes.

Acknowledgements: The authors wish to thank the Santa Rosa County Public School District and the University of West Florida Community Outreach Research and Learning Center for supporting this research partnership

Funding

This work was supported by a contract from the Santa Rosa County School District, Santa Rosa County, Florida, and the Community Outreach, Research, and Learning Center from The University of West Florida, United States of America

Conflict of Interest: none declared.

References

Bush, S. B., Cook, K. L., Ronau, R. W., Rakes, C. R., Mohr-Schroeder, M. J., & Saderholm, J. (2016). A highly structured collaborative STEAM Program: Enacting a professional development framework. *Journal of Research in STEM Education*, *2*(2), 106-125. ISSN: 2149-8504@i-STEM 2015.j-STEM.net.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessments*, *6*(4), 284-290. https://psycnet.apa.org/buy/1995-15835-001

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

Constantino, T. (2018). STEAM by another name: Transdisciplinary practice in art and design education. *Arts Education Policy Review*, *119*(2), 100-106. https://doi.org/10.1080/10632913.2017.1292973.
Duke, N. K. & Halvorsen, A. L. (2017). New study shows the impact of PBL on student achievement. *Edutopia*. https://www.edutopia.org/article/new-study-shows-impact-pbl-student-achievement-nell-duke-anne-lise-halvorsen.
Hackathorn, J. Solomon, E. D., Tennial, R. E., Garcynski, A. M., Blankmeyer, K., Gebhardt, K., & Anthony, J. N., (2010). You get out what you put in: Student engagement affects assessment. Poster presentation. *Best Practices in Assessment* Conference: Atlanta, G. A.
Hackathorn, J. Solomon, E. D., Blankmeyer, K. L., Tennial, R. E., & Garczynski, A. M., (2011). Learning by doing: An empirical study of active teaching techniques. *The Journal for Effective Teaching, 11*(2), 40-54.
Halgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor Quant Methods Psychol,8* (1), 23-24. https://www.ncbi,nlm.nih.gov/pmc/articles/PMC3402032/
Herro, D. Quigley, C. (2017). Exploring teachers' perceptions of STEAM teaching through professional development: Implication for teacher educators. *Professional Development in Education*, *43*(3), 416-438. https://doi.org/10.1080/19415257.2016.1205507.
Hora, M. T., & Ferrare, J. J. (2013). A review of classroom observation techniques in postsecondary settings. (WCER Working Paper No. 2013-1). Retrieved from University of Wisconsin, Madison, Wisconsin. Center for Educational Research.https://wcer.wisc.edu/docs/working_papers/Working_Paper_No_2013_01.pdf
Miller, J. & Knezek, G. (2013). STEAM for Student Engagement. In R. McBride & M. Searson (Eds.), *Proceedings of SITE 2013--Society for Information Technology & Teacher Education International Conference* (pp. 3288-3298). New Orleans, Louisiana, United States: Association for the Advancement of Computing in Education (AACE). Retrieved July 31, 2019 from https://www.learntechlib.org/p/48602.
Nadelson, L. S., Callahan, J. Pike, P., Hay, A., Dance, M., & Pfiester, J. (2013). Teacher STEM perception and preparation: Inquiry-based STEM professional development for elementary teachers. *The Journal of Educational Research, 106*(2), 157-168. https://doi.org/10.1080/00220671.2012.667014
Nadelson, L. S. & Seifert, A. (2017). Integrated STEM defined: Contexts, challenges and the future. *The Journal of Educational Research, 110*(3), 221-223. https://doi.org./10.1080/00220671.2017.1289775.
National Academies Press (2014). STEM integration in K-12Education: Status, Prospects, and an Agenda for Research, Chapter 6, Findings, Recommendations, and Research Agenda. http://www.nap.edu/read/18612/chapter 6/pp.135-152.
Office of the President of the United States (December, 2018). Charting a course for success: America's strategy for STEM education. Committee on STEM Education: National Science & Technology Council. https://www.whitehouse.gov/wp-content/uploads/2018/12/STEM-Education-Strategic-Plan-2018.pdf
Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, *9*(1), 73-90.
Quigley, C. F., Herro, D. & Jamil, F. M. (2017). Developing a conceptual model of STEAM teaching practices. *School Science & Mathematics, 117*(2), 1-12. https://doi.org/10.1111/ssm.12201.
Stuhlman, M. W., Hamre, B. K., Downer, J. T., & Pianta, R. C. (2018). Why should we use classroom observation? Center for Advanced Study of Teaching and Learning. https://curry.virginia.edu/uploads/resourceLibrary/CASTL_practitioner_Part1_single.pdf.

APPENDIX Table 2: Resulting Spearman Rank Order Correlations for Classroom Observation Data with Mathematics Achievement State Standardized Test Scores

Classroom Observed Categories	Numbers and Fractions	Exponents & Equations	Number Operations Base Ten	Measures Data Geometry	Algebra Functions
CP	NS	**NS**	**NS**	**NS**	**NS**
CI	**NS**	**NS**	Rho = .53 $p < .05$	**NS**	**NS**
CTI	**NS**	**NS**	Rho = .42 $p < .10$	**NS**	**NS**
CTP	**NS**	**NS**	Rho = .42 $p < .10$	**NS**	**NS**
CTL	**NS**	**NS**	Rho = .42 $p < .10$	**NS**	Rho =.79 $p < .05$
CDI	**NS**	**NS**	**NS**	**NS**	**NS**
CA	**NS**	**NS**	Rho = .48 $p < .05$	**NS**	**NS**
CTW	**NS**	**NS**	**NS**	**NS**	**NS**
CIS	**NS**	**NS**	**NS**	**NS**	**NS**
Number of Class-rooms	**N= 20**	**N = 20**	**N = 20**	**N = 3**	**N = 6**

APPENDIX Table 2 (continued): Resulting Spearman Rank Order Correlations for Classroom Observation Data with Mathematics Achievement State Standardized Test Scores

Classroom Observed Categories	Algebra Models	Algebra Stat/ Number Systems
CP	**NS**	**NS**
CI	**NS**	**NS**
CTI	**NS**	**NS**
CTP	**NS**	**NS**
CTL	Rho=.75 $p < .05$	Rho=.75 $p < .05$
CDI	**NS**	**NS**
CA	**NS**	**NS**
CTW	**NS**	**NS**
CIS	**NS**	**NS**
Number of Class-rooms	**N = 6**	**N = 6**

APPENDIX Table 3: Resulting Spearman Rank Order Correlations for Classroom Observation Data with Language Arts Achievement State Standardized Test Scores

Classroom Observed Categories	Crafting Structure	Language Editing	Key Ideas & Details	Text-based Writing	Integrating Knowledge & Ideas
CP	NS	**NS**	**NS**	**NS**	**NS**
CI	**NS**	**NS**	**NS**	**NS**	**NS**
CTI	**NS**	**NS**	**NS**	**NS**	**NS**
CTP	**NS**	**NS**	**NS**	**NS**	**NS**
CTL	**NS**	**NS**	**NS**	**NS**	**NS**
CDI	**NS**	**NS**	**NS**	**NS**	**NS**
CA	**NS**	**NS**	NS	NS	**NS**
CTW	**NS**	**NS**	**NS**	**NS**	**NS**
CIS	**NS**	**NS**	**NS**	**Rho=.37 $p < .10$**	**NS**
Number of Classrooms	**N= 20**	**N = 20**	**N= 20**	**N = 20**	**N = 20**