

The impact of technology use and teacher professional development on U.S. national assessment of educational progress (NAEP) mathematics achievement

Byron Havard¹  · Giang-Nguyen Nguyen² · Barbara Otto³

Received: 9 April 2017 / Accepted: 15 February 2018 / Published online: 9 March 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract The purpose of this study was to determine the impact of technology use and teacher professional development on students' mathematics academic achievement. The U.S. Department of Education National Assessment of Educational Progress (NAEP) published results for mathematics assessments for Grade 4 from the years 2005–2015 served as the dependent variable. Specific items related to technology use and professional development selected from both student and mathematics teacher questionnaires served as the independent variables. The Technological Pedagogical and Content Knowledge (TPACK) was used as a framework to guide this research. Data analyses revealed significant differences across multiple variables and multiple years.

Keywords Technology use · Computers · Calculators · Mathematics · Professional development

✉ Byron Havard
bhavard@uwf.edu

Giang-Nguyen Nguyen
nguyen@uwf.edu

Barbara Otto
ottob@uni-landau.de

- ¹ Instructional, Workforce and Applied Technology, University of West Florida, 70/107, Pensacola, FL 32514, USA
- ² Teacher Education and Educational Leadership, University of West Florida, 85/195, Pensacola, FL 32514, USA
- ³ Department of Psychology, University of Koblenz-Landau, Campus Landau Fortstraße 7, D-76829 Landau in der Pfalz, Germany

1 Introduction

The U.S. Department of Education National Assessment of Educational Progress (NAEP) provides a common measure of student comprehension across a wide variety of subject areas. A sample of students in Grades 4, 8, and 12 participated in NAEP assessments representing the national population in a variety of subject areas. According to the NAEP, 60% of students performed below proficient in mathematics (42% at the basic level and 18% below the basic level) in 2015 even though there were some improvements in mathematics performance compared to previous years (NCES 2016a).

Several meta-analyses have revealed positive effects of technology on academic achievement. In an analysis of 46 studies, including 85 independent findings and encompassing 36,793 students, computer technology revealed a positive effect ($ES = +0.71$) on mathematics achievement (Li and Ma 2010). Another analysis reviewing 58 studies noted positive effects of technology use on both cognitive ($ES = +0.42$) and affective ($ES = +0.18$) learning outcomes (Lee et al. 2013). A larger scale analysis of 74 studies revealed an overall positive effect ($ES = +0.15$) in classes that used technologies to enhance mathematics instruction as compared to traditional teaching methods (Cheung and Slavin 2013). These empirical findings point to the immense potential effective technology use could have on mathematics education. While this present study offers results that are consistent with the positive effects revealed in the meta-analyses, we propose that missing elements related to teacher professional development exist.

Many efforts have been taken to understand why students are not performing well in mathematics (NCES 2015a). Specifically, one possible reason for students' low performance in mathematics might be deficits in mathematics teacher professional development regarding the integration of technology into teaching and learning (Heid and Blume 2008). Research in this area is lacking, and because of that, some technology integration programs have not been successful in terms of student learning due to an absence of best practices (Means 2010).

Technology integration in K-12 education has been encouraged by national, state, and regional accreditation agencies through initiatives including the National Education Technology Plan and the Enhancing Education through Technology (Ed-Tech) State Program among others (Horn and Mackey 2011; Kennedy 2010). School districts provide mathematics teacher professional development with the intent to foster and cultivate various forms of teacher knowledge which in turn may improve student learning. However, research lacks in investigating the specific linkage between mathematics teachers' professional development and their students' performance in mathematics. Therefore, examining the impact of professional development over time is recommended (Dash et al. 2012). This present study specifically examines the influence of technology use and professional development on mathematics achievement over time.

1.1 Research purpose

The purpose of the study was to determine the impact of technology use and teacher professional development on students' mathematics academic achievement. The Technological Pedagogical Content Knowledge (Mishra and Koehler 2006) was used as a

framework to guide the study. To address the research questions, we used the NAEP published results for mathematics assessments for Grade 4 from the years 2005–2015. Grade 4 students' NAEP mathematics assessment average scale scores were used in the analysis as the dependent variable; specific items related to technology use and professional development selected from both student and mathematics teacher questionnaires served as the independent variables.

1.2 Research rationale

Various professional development research efforts focused on the improvement of student learning have produced mixed results (Hill et al. 2013). However, information derived from these studies contributes to a stronger understanding of the necessary characteristics of successful professional development. In that respect, we want to examine the influence of technology use and professional development on students' mathematics achievement. If such use of technology and professional development does not improve student learning, it will be important to explore further why it is not working. The results could provide some directions for future professional development that supports teachers and students as they make use of technologies for the teaching and learning of mathematics. This paper is organized into the following sections: (a) Background, (b) Conceptual framework, (c) Methodology, (d) Results, (e) Discussion, (f) Implications, and (g) Conclusion.

2 Background

Several meta-analyses have been conducted using previous research on media effects on mathematics achievement, educational technology, and the relationship between the two (Campuzano et al. 2009; Cheung and Slavin 2013; Dynarski et al. 2007; Li and Ma 2010). Cheung and Slavin (2013), analyzed 74 studies with a total of 56,886 K-12 students. In this large-scale meta-analysis, researchers found that application of educational technology defined as “a variety of technology-based programs or applications that help deliver learning materials and support learning process in K-12 classrooms to improve academic learning goals” (p. 90), had a positive effect on K-12 students' mathematics achievement. Previous research studies' findings (Campuzano et al. 2009; Dynarski et al. 2007) partially aligned with Cheung and Slavin, but the difference is largely based on sample size differences. In particular, smaller sample sized studies showed about twice the effect sizes as their larger counterparts. It was also found that, though all K-12 students benefit from educational technology applications, secondary instruction does not benefit from technology as much as primary instruction (Cheung and Slavin 2013). Another interesting finding concerns the amount of time students actually use the technology. It was assumed that time limitations were because of failure to implement properly. However, the authors suggest the issue may be because of teachers not recognizing the value of the tools, thus not using them to their full potential. Studies and reviews of this nature are limited by the speed at which technology is evolving, which prevents long-term studies. While this study review showed the educational technology applications only

have a moderate effect on mathematics achievement, it also shows that the process is not stagnant and educators are constantly learning how to better utilize tools to make this effect more significant.

Technology integration has also been shown to have a potentially negative impact on academic achievement, especially in terms of level of use and how the technology is integrated. This finding is based on Programme for International Student Assessment data from 34 Organisation for Economic Co-Operation and Development (OECD) countries.

Students who use computers moderately at school tend to have somewhat better learning outcomes than students who use computers rarely. But students who use computers very frequently at school do a lot worse in most learning outcomes, even after accounting for social background and student demographics. Extensive use of technology was found to be detrimental to academic performance. (OECD 2015 p. 3)

Rather than access to technology, the focus must address how technology is being used (OECD 2015). Roschelle et al. (2010) developed an approach that combines representational software, curriculum, and teacher professional development referred to as SimCalc for Grades 7 and 8 math students. Two randomized experiments and one quasi-experiment revealed positive student-level effect sizes, .63, .50, and .56., of SimCalc on mathematics achievement scores. The emphasis of this approach, based on design research, was the integration of the software, curriculum, and professional development. These three components may prove to be a potentially successful combination based on a strong foundation of teachers' mathematics content knowledge and pedagogical knowledge (Shulman 1986).

In the teaching field, as with many fields, research and practice go hand-in-hand. Even though practices are usually research-based, new research on teaching practices is continuously conducted (Heid and Blume 2008). In response to the need for professional development, some efforts have been made to offer more opportunities such as online professional development and webinars. Massive Open Online Courses (MOOCs), more frequently used in higher education, are now offered for K-12 teacher professional development (Briggs and Crompton 2016). Dash et al. (2012) explored the way online professional development for mathematics instruction impacts pedagogy, knowledge of content and strategy practices, and in turn the achievement level of their students. Dash et al. (2012), pointed out that “true effects of professional development on student achievement cannot be ascertained without first considering teachers' opportunity to implement the material” (p. 22) suggesting that professional development needs to be accompanied with some forms of support.

In light of the aforementioned limitation, it is proposed that further research is needed. Dash et al. (2012), recommend following up with those receiving professional development to measure effects over time. Also, the researchers suggest studies of support systems following professional development in order to determine how best to support teachers' implementation of their newfound knowledge. Finally, an expansion of the study of online professional development to include face-to-face courses is mentioned in order to compare the effectiveness of mode.

3 Conceptual framework

The Technological Pedagogical and Content Knowledge (TPACK) was used as a framework to guide this research study. The evolution of this framework, relationship to the present study, and relationship to the variables will be described in this section.

3.1 Teacher's professional competence and technological pedagogical content knowledge

Baumert and Kunter (2006) investigated teachers' professional competence focusing on knowledge as the key element of competence. Though teacher knowledge is a multifaceted concept, the researchers focused on content knowledge, pedagogical content knowledge, and general pedagogical/psychological knowledge. To empirically test these aspects, Baumert and Kunter sought to measure teachers' content knowledge (specifically mathematics), organization and preparation, and intuitive responses. They indicated the necessity to approach content knowledge, not only as an understanding of the material covered but also as an understanding of the approaches to expressing that knowledge. Pedagogical content knowledge was defined as knowledge of the diagnostic value of the material, of student cognitive processes, and of varied instruction/demonstration. Teachers must have a broad understanding of the content in order to shape comprehension, use appropriate strategies, and express themselves with those not possessing the same level of knowledge (Baumert and Kunter 2006, 2013).

3.2 Technological pedagogical content knowledge

Teaching mathematics requires a teacher to possess various forms of knowledge in addition to the content knowledge of mathematics (Baumert and Kunter 2006; Ernest 1989; Shulman 1986). While content knowledge (CK) and pedagogical knowledge (PK) are fundamental for effective teaching, the knowledge of knowing how students learn mathematical content, referred to as pedagogical content knowledge (PCK), is even more important (Baumert and Kunter 2013; Kunter et al. 2013). In the PCK framework, teachers must have an understanding of how students learn mathematics to teach mathematics content. Kunter et al. (2013) found that PCK was a stronger predictor of instructional quality and student progress compared to content knowledge (CK).

In this pedagogical context, technology may provide many opportunities to turn abstract mathematical content knowledge into hands-on mathematics. Technological pedagogical knowledge (TPK), knowing how to integrate technology into teaching and learning mathematics, is an important component of teacher knowledge (Handal et al. 2013). However, it is not enough for teachers to have the knowledge of technology as tools but they must have the knowledge of how to incorporate various technologies to enhance students learning experience in learning the content. Mishra and Koehler (2006) stated that there must be some “connections, interactions, affordances, and constraints” (p. 1025) between and among the knowledge of technology, pedagogy, and content of which they developed into a framework. The Technological Pedagogical Content Knowledge (TPCK) framework provides a foundation for research and practice regarding teacher knowledge and technology integration (Koehler et al. 2007; Mishra and Koehler 2006; Yurdakul et al. 2012).

This framework, now referred to as TPACK, has emerged as a viable framework for mathematics teacher professional development (Polly 2011). In this framework, three types of knowledge are distinguished. Pedagogical Content Knowledge (PCK) serves as the basis for TPACK, describing the dynamic interplay between teachers' knowledge of content and teaching strategies for useful representation in a learning environment (Mishra and Koehler 2006; Shulman 1986). The increase in technology integration in educational environments stimulated the need for the unifying TPACK framework, providing researchers an organizational scheme regarding the interplay among teachers' technology, pedagogy, and content knowledge. Technological Content Knowledge (TCK) includes how various representations of content may be used in learning environments such as the use of Geometer's Sketchpad for teaching geometry (Mishra and Koehler 2006). Finally, Technological Pedagogical Knowledge (TPK) relates to the technology-enabled strategies used based on the "understanding that a range of tools exists for a particular task, the ability to choose a tool based on its fitness, strategies for using the tool's affordances, and knowledge of pedagogical strategies and the ability to apply those strategies for use of technologies" (Mishra and Koehler 2006 p. 1028). The integration of all three components of the framework, namely technology, pedagogy, and content, provides a professional development learning framework regarding the many facets of emerging technologies, pedagogical strategies, and concept representations for learning (Koehler and Mishra 2009; Koehler et al. 2007; Mishra and Koehler 2006; Yurdakul et al. 2012).

3.3 Technological pedagogical content knowledge and the present study

Studying the influence of technology use and professional development on student mathematics achievement, and whether the influence is positive or negative, informs future professional development practice. The results could provide direction for professional development that supports teachers and students as they make use of technologies for the teaching and learning of mathematics. We used the TPACK framework in this study because teacher professional development on the use of technologies such as computers and calculators should influence TPK, as this knowledge is necessary for effective technology integration to positively impact mathematics achievement (Mishra and Koehler 2006).

3.4 Data variables

NAEP survey questionnaires are "developed and reviewed by survey experts, education researchers, teachers, and statisticians to ensure that the information collected is relevant and valid in helping policymakers, researchers, educators, and the public understand student achievement results." (NCES 2017 para. 4). Variables were purposefully selected based on the TPACK framework. An example of each data variable is provided below:

Computer Variable Example: In your fourth-grade math class this year, how often do your students use a computer or other technological resources to extend mathematics learning with enrichment activities on the computer? (a) Never or hardly ever, (b) Once or twice a month, (c) Once or twice a week, (d) Every day or almost every day.

Calculator Variable Example: To what extent are students permitted to use calculators during mathematics lessons? (a) Unrestricted use, (b) Restricted use, (c) Calculators not permitted.

Professional Development Variable Example: Consider all of the professional development activities you participated in during the last two years. To what extent did you learn about each of the following topics? Effective use of calculators in mathematics instruction. (a) Not at all, (b) Small extent, (c) Moderate extent, (d) Large extent.

4 Methodology

In this section, research questions are presented, data source information is provided, instruments are described, and a discussion of how data were analyzed is included. The term ‘educational technology’ was used as described in previous research by Cheung and Slavin (2013) and was defined in the background section of this present study. An example question on application of educational technology: In your fourth-grade math class this year, how often do your students use a computer or other technological resources to (a) practice or review mathematics, (b) extend mathematics learning, (c) research mathematics topics, and (d) play mathematics games. Details of items used will be discussed the instrument sections.

4.1 Research questions

A well-balanced integration of technology in mathematics instruction can be assumed to foster students’ learning and achievement (Cheung and Slavin 2013; Li and Ma 2010). However, a focus on effective strategies for technology integration is often neglected due to little emphasis on teacher professional development (Means 2010). Accordingly, in this study we wanted to explore the impact of technology use and professional development on Grade 4 students’ NAEP mathematics assessment average scale scores. In this context the following three research questions served to guide this study:

- 1) Is there a statistically significant difference in student mathematics achievement on the NAEP Grade 4 assessment average scale scores between students who used computers in mathematics classes compared to those who were not permitted to use computers?
- 2) Is there a statistically significant difference in student mathematics achievement on the NAEP Grade 4 assessment average scale scores between students who used calculators in mathematics classes compared to those who were not permitted to use calculators?
- 3) Is there a statistically significant difference in student mathematics achievement on the NAEP Grade 4 assessment average scale scores between students with teachers that participated in professional development of integrating technology into mathematics classes compared to those that did not participate in professional development?

4.2 Data source

The NAEP published mathematics assessment results for Grade 4 from the years 2015, 2013, 2011, 2009, 2007, and 2005 were used as the dataset for this study (NCES 2016b). Students are selected for participation in the NAEP mathematics assessment through a probability sample design of Grade 4 students, approximately 9 years of age, in the United States from both public and private schools (NCES 2011). Schools in each state are grouped by strata based on public school characteristics including government-defined geographic area (jurisdictions), minority enrollment, scores on state-based assessments, and median income of the school location. Student sampling for the assessment from each jurisdiction includes the selection of approximately 100 schools and approximately 60 students from each school (NCES 2011). Teachers are sampled from schools with at least one student participating in the assessment. To ensure unbiased samples, participation rates of 70% or higher are required for sampled schools. Non-response bias analysis is conducted where participation rates exceed 70% but fall below 85%. If bias is detected additional schools are substituted for originally sampled schools. Results are weighted and adjusted for any disproportions to reflect the larger representative population (NCES 2015b). Table 1 provides information about the population and the sample for the analyses in the present study.

4.3 Instruments

4.3.1 Student performance based on NAEP mathematics assessment scores

The results of the present study based on the assessment of fourth graders' mathematical competencies measured by the NAEP Mathematics Assessment as well as teacher and student reports about technology integration and professional development. The assessment items relate to five content areas with items distributed based on the following percentages (National Assessment Governing Board [NAGB] 2012): Number Properties and Operations (40%); (b) Measurement (20%); Geometry (15%); Data Analysis, Statistics, and Probability (10%); and Algebra (15%). Item formats are distributed evenly between multiple choice (50%) and both short and extended constructed response (50%). Item complexity, the cognitive demands of the item, include low (25%), moderate (50%), and high complexity (25%). "Items at the low level of complexity, for example, may ask a student to recall a property. At the moderate level, an item may ask the student to make a connection between two properties; at the high level, an item may ask a

Table 1 Student population and sample

Year	Grade 4 Population	Sample
2015	3,939,000	142,600
2013	3,895,000	189,660
2011	3,945,000	214,200
2009	3,824,000	173,300
2007	3,798,000	204,000
2005	4,174,000	178,000

Source: National Center for Education Statistics (2006, 2007, 2010, 2012, 2013, 2016)

student to analyze the assumptions made in a mathematical model” (NAGB 2012 p. 37). Multiple forms of the assessment are created through matrix sampling; students do not receive the same items. This permits a broader range of content objectives to be assessed in a shorter amount of time; namely two 25-min assessment periods (NAGB 2007, 2012). Multiple choice items are scanned and scored. Extensive training is provided for scorers of the short and extended-response items. Scorers are selected based on prior experience and qualifying tests. Even though the authors did not provide the exact reliability information, scoring consistency was maintained through consistent scorer reliability checks and assessment of scorer decision-making (NCES 2014). Additionally, non-cognitive items were also included in the student questionnaires.

4.3.2 Non-cognitive items on teacher and student questionnaires

Questionnaires are used to gather non-cognitive assessment information. Student questionnaires consist of two sections relating to (a) Demographic Characteristics and (b) Classroom Experiences and Educational Support. Each section requires approximately five minutes for students to complete. Teacher questionnaires include two sections: (a) Background, Education, and Training and (b) Classroom Organization and Instruction (NCES 2015c). In order to investigate the influence of technology use and professional development over time, the following aspects were measured: computer use, calculator use, and teacher professional development.

4.3.3 Computer use variables on teacher questionnaires

Four teacher-reported variables were used to assess the impact of students’ use of computers on NAEP mathematics average scale scores. Teachers responded to the following question: In your fourth-grade math class this year, how often do your students use a computer or other technological resources to (a) practice or review mathematics, (b) extend mathematics learning, (c) research mathematics topics, and (d) play mathematics games. Potential responses included the following values: (a) never or hardly ever, (b) once or twice a month, (c) once or twice a week, (d) every day or almost every day. Data were gathered at four separate time periods: 2015, 2013, 2011, and 2009.

4.3.4 Calculator use variables on teacher and student questionnaires

Two teacher-reported variables and two student-reported variables were used to assess the impact of calculator use on NAEP mathematics average scale scores. These variables were further differentiated between the use of a calculator for mathematics lessons and the use of a calculator for mathematics tests. Data were gathered at six separate time periods: 2015, 2013, 2011, 2009, 2007, and 2005.

4.3.5 Teacher professional development variables on teacher questionnaires

Three teacher-reported variables were used to assess the impact of teacher professional development on NAEP mathematics average scale scores. Teachers

responded to the following question: Consider all of the professional development activities you participated in during the last two years. To what extent did you learn about each of the following topics? The topics selected by the researchers included professional development related to (a) instructional methods for teaching mathematics, (b) effective use of calculators in mathematics instruction, and (c) use of computers or other technology in mathematics instruction. Potential responses included the following values: (a) not at all, (b) small extent, (c) moderate extent, (d) large extent. Data were gathered at six separate time periods: 2015, 2013, 2011, 2009, 2007, and 2005. For the purpose of this study it is important to note that only teachers' extent of participation in professional development is measured; whether teachers' adequately implemented knowledge or skills learned as a result of the professional development or measures of the quality or effectiveness of the professional development are beyond the scope of this study.

4.4 Data analysis

Data were analyzed through analyses of variance (ANOVA) using the NAEP Data Explorer (NCES 2016a). While the authors proposed to analyze data through multi-variate analysis of variance (MANOVA) to reduce the chance of Type 1 error, MANOVA was not offered through the NAEP Data Explorer because NAEP does not allow researchers to export raw data for analysis. This may be viewed as a limitation of this study. As the result, all data were analyzed through NAEP Data Explorer.

5 Results

The results of the data analysis are presented here in three sections based on the three research questions. The three sections are comprised of (a) computer use in mathematics teaching and learning, (b) calculator use in mathematics teaching and learning, and (c) teacher professional development in mathematics teaching and learning.

5.1 Computer use in mathematics teaching and learning

Four variables were used to examine computer use in mathematics teaching and learning. The results of the data analysis are presented in four sub-sections relating to those four variables comprised of computer use (a) for practice or review of mathematics, (b) to extend mathematics learning, (c) to research mathematics topics, and (d) to play mathematics games.

5.1.1 Computer use for practice or review of mathematics

Data analysis regarding students' use of computers to practice or review mathematics topics *once or twice a month* revealed a statistically significant ($p < .01$) difference (higher scores) in Grade 4 assessment scale scores across all four data collection periods compared to *never or hardly ever* using the computer for this

purpose. Student computer use to practice or review mathematics topics *once or twice a week* generated statistically significant ($p < .01$) differences (higher score) in 2015 and (lower score) in 2011. Use *every day or almost every day* resulted in statistically significant ($p < .01$) differences (higher score) in 2015 and (lower score) in 2009 (see Table 2).

5.1.2 Computer use to extend mathematics learning

Student computer use to extend mathematics learning *once or twice a month* showed statistically significant ($p < .01$) differences (higher) in Grade 4 assessment scale scores across all four data collection periods compared to *never or hardly ever* using the computer for this purpose. Student computer use to extend mathematics learning *once or twice a week* differed significantly ($p < .01$) with a higher score in 2015 and lower scores in 2011 and 2009, and use *every day or almost every day* differed significantly ($p < .01$) resulting in lower scores in 2013 and 2009.

5.1.3 Computer use to research mathematics topics

Student computer use to research mathematics topics *once or twice a month* differed significantly ($p < .01$) with a higher Grade 4 assessment scale score in 2013 compared to *never or hardly ever* using the computer to research mathematics topics. Student

Table 2 Average scale scores for grade 4 mathematics by students' use of computer variables

Use of computer variables	Year	Never or hardly ever	Once or twice a month	Once or twice a week	Every day or almost every day
Practice or review mathematics	2015	239 (0.7)	242 (0.5)**+	241 (0.4)**+	241 (0.7) **+
	2013	241 (0.7)	245 (0.5)**+	241 (0.4)	240 (0.6)
	2011	241 (0.5)	244 (0.4)**+	239 (0.4)**-	240 (0.9)
	2009	240 (0.5)	242 (0.4)**+	239 (0.4)	237 (0.6)**-
Extend mathematics learning	2015	239 (0.6)	242 (0.4)**+	241 (0.4)**+	241 (0.9)
	2013	241 (0.5)	244 (0.5)**+	241 (0.4)	239 (0.8)**-
	2011	241 (0.4)	243 (0.4)**+	239 (0.5)**-	239 (1.2)
	2009	240 (0.4)	242 (0.4)**+	238 (0.5)**-	236 (0.8)**-
Research mathematics topics	2015	≠ ±	≠ ±	≠ ±	≠ ±
	2013	242 (0.2)	244 (0.6)**+	240 (0.8)	236 (1.6)**-
	2011	241 (0.2)	242 (0.5)	238 (0.8)**-	238 (2.1)
	2009	240 (0.3)	240 (0.5)	237 (0.9)**-	236 (1.3)**-
Play mathematics games	2015	239 (0.7)	242 (0.4)**+	241 (0.4)	241 (0.9)
	2013	241 (0.7)	244 (0.4)**+	241 (0.4)	240 (0.8)
	2011	241 (0.5)	243 (0.4)**+	240 (0.4)	240 (0.8)
	2009	240 (0.6)	242 (0.4)**+	239 (0.4)	236 (0.8)**-

Standard error shown in parentheses, - indicates significantly lower score, + indicates significantly higher score, ≠ indicates not available, ± indicates not applicable

* $p < .05$, ** $p < .01$

computer use to research mathematics topics *once or twice a week* differed significantly ($p < .01$) revealing lower scores in 2011 and 2009, and use *every day or almost every day* differed significantly ($p < .01$) with a lower score in 2009. The 2015 data were not available for this variable.

5.1.4 Computer use to play mathematics games

Student computer use to play mathematics games *once or twice a month* revealed statistically significant ($p < .01$) differences generating higher Grade 4 assessment scale scores in 2015, 2013, 2011, and 2009 compared to *never or hardly ever* using the computer to play mathematics games. Student computer use to play mathematics games *every day or almost every day* produced a statistically significant ($p < .01$) difference (lower) score in 2009. For example, Table 2 illustrates, in the year 2013, we compared (a) *never or hardly ever* to *once or twice a month*, yielding a significant different ($p < .01$), (b) *never or hardly ever* compared to *once or twice a week*, did not yield a significant difference, and (c) *never or hardly ever* compared to *every day or almost every day*, did not yield a significant difference. In each of the analyses in Table 2 we compared *never or hardly ever* to the other levels of use.

5.2 Calculator use in mathematics teaching and learning

Four variables were used to examine calculator use in mathematics teaching and learning. The results of the data analysis are presented in two subsections, each with two variables, comprised of calculator use for mathematics lessons and calculator use for mathematics tests.

5.2.1 Calculator use for mathematics lessons

With regard to calculator use for mathematics lessons, teachers responded to the following question: To what extent are students permitted to use calculators during mathematics lessons? Potential responses included the following values: (a) not permitted, (b) restricted use, and (c) unrestricted use. *Restricted use* of a calculator during mathematics lessons revealed statistically significant ($p < .01$) difference (higher) in Grade 4 assessment scale scores in all six data collection periods when compared to calculators *not permitted*. *Unrestricted use* yielded significantly ($p < .01$) different (higher) Grade 4 assessment scale scores in 2013. Students responded to the following question regarding calculator use for mathematics lessons: How often do you use a calculator [during mathematics lessons]? Potential responses included the following values: (a) never or hardly ever, (b) once in a while, and (c) almost every day. Students' reported use of calculators *once in a while* produced significantly ($p < .01$) different (lower) Grade 4 assessment scale scores in 2015 and 2013 compared to *never or hardly ever* using a calculator for math lessons. Calculator use *almost every day* for math lessons yielded statistically significant ($p < .01$) differences (lower) in Grade 4 assessment scale scores in five data collection periods (see Table 3). Student reported data for 2007 regarding this variable are not available.

Table 3 Average scale scores for grade 4 mathematics by teacher and student reported students' use of calculator for math lessons

Use calculator for math lessons	Year	Not permitted	Restricted use	Unrestricted use
Teacher reported	2015	238 (1.8)	243 (0.4)**+	238 (0.5)
	2013	236 (1.2)	244 (0.3)**+	239 (0.3)**+
	2011	236 (0.9)	243 (0.3)**+	237 (0.4)
	2009	235 (1.0)	242 (0.2)**+	237 (0.5)
	2007	235 (0.9)	242 (0.2)**+	236 (0.4)
	2005	233 (0.9)	239 (0.2)**+	235 (0.4)
	Student reported	Year	Never or hardly ever	Once in a while
2015		243 (0.3)	241 (0.3)**-	221 (0.5)**-
2013		245 (0.3)	242 (0.3)**-	224 (0.6)**-
2011		242 (0.3)	242 (0.2)	226 (0.5)**-
2009		241 (0.4)	241 (0.2)	225 (0.5)**-
2007		≠ ±	≠ ±	≠ ±
2005		239 (0.2)	239 (0.1)	226 (0.4)**-

Standard error shown in parentheses, – indicates significantly lower score, + indicates significantly higher score, ≠ indicates not available, ± indicates not applicable

* $p < .05$, ** $p < .01$

5.2.2 Calculator use for mathematics tests

The impact of calculator use for mathematics tests was measured through the following teacher and student questions. Teachers were asked: When you give students a mathematics test or quiz, how often do they use a calculator? Students were asked a similar question: When you take a math test or quiz, how often do you use a calculator? Potential responses for both the teacher and student questions included the following values: (a) never, (b) sometimes, and (c) always. Use of a calculator for math tests *sometimes*, as reported by teachers, revealed a statistically significant ($p < .01$) difference (higher) in Grade 4 assessment scale scores for five of the six data collection periods compared to *never* using a calculator. Use of a calculator for mathematics tests *always*, as reported by teachers, produced a statistically significant ($p < .01$) difference (lower) in Grade 4 assessment scale scores for all six data collection periods. Use of a calculator for math tests both *sometimes* and *always*, as reported by students, generated significantly ($p < .01$) different (lower) Grade 4 assessment scale scores for all six data collection periods compared to *never* using a calculator for tests (see Table 4).

5.3 Teacher professional development in mathematics teaching and learning

Three variables were used to examine teacher professional development in mathematics teaching and learning. The results of the data analysis are presented in three subsections comprised of teacher professional development on (a) instructional methods for teaching mathematics, (b) effective use of calculators in mathematics instruction, and (c) the use of computers or other technology in mathematics instruction.

Table 4 Average scale scores for grade 4 mathematics by teacher and student reported students' use of calculator for math tests

Use calculator for math tests	Year	Never	Sometimes	Always
Teacher reported	2015	241 (0.3)	242 (0.6)	226 (3.2)**-
	2013	241 (0.3)	243 (0.4)**+	232 (2.1)**-
	2011	240 (0.3)	243 (0.4)**+	232 (1.4)**-
	2009	240 (0.3)	241 (0.3)**+	231 (1.7)**-
	2007	240 (0.2)	241 (0.3)**+	231 (1.7)**-
	2005	238 (0.2)	240 (0.4)**+	229 (2.9)**-
Student reported	2015	243 (0.3)	235 (0.4) **-	211 (1.3)**-
	2013	245 (0.3)	237 (0.3)**-	211 (0.9)**-
	2011	243 (0.3)	237 (0.2)**-	211 (0.8)**-
	2009	242 (0.3)	237 (0.3)**-	208 (1.1)**-
	2007	243 (0.2)	237 (0.2)**-	209 (1.1)**-
	2005	240 (0.2)	235 (0.2)**-	207 (1.0)**-

Standard error shown in parentheses, – indicates significantly lower score, + indicates significantly higher score

* $p < .05$, ** $p < .01$

5.3.1 Teacher professional development on instructional methods for teaching mathematics

Data analysis regarding teacher professional development on instructional methods for teaching mathematics did not reveal any statistically significant differences in Grade 4 assessment scale scores across all six data collection periods when compared to no professional development. Teacher professional development in instructional methods for teaching mathematics in *small*, *moderate*, and *large* extents revealed a statistically significant difference ($p < .01$) in scores (lower) in 2005 compared to no professional development (see Table 5).

5.3.2 Teacher professional development on effective use of calculators in mathematics instruction

Teacher professional development on the effective use of calculators in mathematics instruction in a *small extent* produced a statistically significant difference ($p < .01$) in scores (lower) in 2009. A *moderate extent* differed significantly ($p < .01$) with lower scores in 2013, 2011, 2009, and 2005, and a *large extent* differed significantly ($p < .01$) with lower scores in 2011, 2009, and 2005 compared to no professional development.

5.3.3 Teacher professional development on the use of computers or other technology in mathematics instruction

Teacher professional development on the use of computers or other technology in mathematics instruction revealed statistically significant differences ($p < .01$) in Grade

Table 5 Average scale scores for grade 4 mathematics by teacher professional development variables

Teacher professional development variables	Year	Not at all	Small extent	Moderate extent	Large extent
Instructional methods for teaching mathematics	2015	241 (0.9)	240 (0.4)	241 (0.4)	241 (0.5)
	2013	241 (0.8)	241 (0.5)	242 (0.3)	242 (0.4)
	2011	241 (0.6)	241 (0.4)	240 (0.3)	241 (0.6)
	2009	240 (0.6)	240 (0.4)	239 (0.4)	240 (0.5)
	2007	240 (0.5)	241 (0.3)	239 (0.3)	240 (0.4)
	2005	240 (0.4)	238 (0.3)**-	238 (0.3)**-	238 (0.4)**-
Effective use of calculators in mathematics instruction	2015	241 (0.3)	241 (0.5)	240 (0.9)	239 (1.3)
	2013	242 (0.3)	242 (0.5)	240 (0.6)**-	240 (1.2)
	2011	241 (0.3)	241 (0.4)	239 (0.6)**-	237 (1.1)**-
	2009	241 (0.3)	239 (0.4)**-	238 (0.5)**-	235 (1.2)**-
	2007	240 (0.3)	240 (0.3)	240 (0.4)	237 (1.3)
	2005	239 (0.2)	238 (0.3)	237 (0.5)**-	234 (1.0)**-
Use of computers or other technology in mathematics instruction	2015	240 (0.5)	241 (0.4)	241 (0.6)	241 (0.9)
	2013	241 (0.5)	242 (0.3)**+	242 (0.5)	241 (0.8)
	2011	241 (0.5)	241 (0.3)	241 (0.4)	241 (0.8)
	2009	240 (0.4)	240 (0.4)	240 (0.4)	239 (0.8)
	2007	239 (0.4)	240 (0.3)	240 (0.4)	240 (0.7)
	2005	238 (0.3)	238 (0.2)	238 (0.4)	237 (0.6)**-

Standard error shown in parentheses, – indicates significantly lower score, + indicates significantly higher score
 * $p < .05$, ** $p < .01$

4 assessment scale scores in two of the five data collection time periods. A *small extent* of teacher professional development in use of computers or other technology in mathematics instruction produced a statistically significant ($p < .01$) difference with a higher Grade 4 assessment scale score in 2013 when compared to no professional development. A *large extent* of professional development revealed a statistically significant difference ($p < .01$) in a lower score in 2005.

6 Discussion

6.1 Computer use in mathematics teaching and learning

The first research question was phrased as follow: *Is there a statistically significant difference in student mathematics achievement on the Grade 4 NAEP assessment average scale scores between students who used computers in mathematics classes compared to those who were not permitted to use computers?* The four computer use variables revealed significant differences in NAEP assessment average scale scores across all four data collection periods. Computer use *once or twice a month* showed statistically significant ($p < .01$) differences (higher) in Grade 4 assessment average scale scores across all four variables for all years compared to *never or hardly ever* using a

computer, except 2011 and 2009 for the *research mathematics topics* variable. Computer use once or twice a week and every day or almost every day showed statistically significant ($p < .01$) differences (lower) in Grade 4 assessment average scale scores across all four variables for roughly one-third of the data collection periods. However, Tienken and Wilson (2007) found the use of computer-assisted instruction enhanced basic mathematical computational skills for students that accessed websites that featured mathematics drill and practice sessions twice a week for 20 weeks. These students also used the computer to present mathematical concepts they learned from the drill and practice sessions. This type of integration, where learned concepts are further reinforced, may enhance retention rather than isolated drill and practice sessions. Kim and Chang's (2010) study of NAEP 2005 Grade 4 students revealed significantly lower math achievement scores for English speaking students who played math games every day compared to those that never played math games. Carr (2012) found no significant difference in Grade 5 mathematics learning outcomes, where the treatment group used iPad games daily for nine months and the control group did not. However, Bakker et al. (2015) found that second and third graders that played mathematics games at home followed by debrief sessions at school positively affected students' multiplicative operations skills and insight. The results and supporting research appear to suggest limiting computer use and integrating both home and school use into mathematics instruction.

6.2 Calculator use in mathematics teaching and learning

The second research question was phrased as follows: *Is there a statistically significant difference in student mathematics achievement on the NAEP Grade 4 assessment average scale scores between students who used calculators in mathematics classes compared to those who were not permitted to use calculators?* Teacher-reported restricted calculator use for mathematics lessons revealed statistically significant ($p < .01$) differences (higher) in Grade 4 assessment average scale scores across all six data collection periods while student-reported calculator use *almost every day* showed statistically significant ($p < .01$) differences (lower scores). Based on this finding calculator use should be managed by the teacher to be used most effectively. Teacher-reported restricted calculator use *sometimes* for mathematics tests revealed statistically significant ($p < .01$) differences (higher) in Grade 4 assessment average scale scores across five of the six data collection periods. Statistically significant ($p < .01$) differences (lower scores) were revealed for teacher-reported always using calculators for tests and student reported sometimes and always using calculators for tests. A considerable number of research studies were conducted in the 1970s regarding calculator use during mathematics instruction and assessments. Roberts (1980) reviewed 34 empirical studies concerning the effects of calculator use during testing among U.S. elementary, secondary, and university level students. The design of a majority of the studies included a pretest, a posttest, an experimental group (used calculators), and a control group (no calculators). The majority of studies cited no significant difference between experimental and control groups (Roberts 1980). Walcott and Stickles' (2012) review of research on the use of calculators in the Grade 4 revealed a neutral effect for basic math skills and a positive effect for fostering higher-order problem-solving skills in the eighth grade. Calculators should be used to promote problem-solving since their use influences and encourages decision making

(Thompson and Sproule 2000; Van de Walle et al. 2007). Bouck et al. (2013) found that calculator use by sixth and seventh-grade math students, both with and without disabilities, resulted in more problems answered correctly. Calculators have been used effectively by students in primary grades (Polly 2008) so the issue may be related to teacher professional development or effective integration of calculator use in mathematics curriculum and instruction.

6.3 Teacher professional development in mathematics teaching and learning

The third research question was phrased as follows: *Is there a statistically significant difference in student mathematics achievement on the NAEP assessment average scale scores between students with teachers that participated in professional development of integrating technology into mathematics classes compared to those that did not participate in professional development?* Teacher professional development in instructional methods for teaching mathematics in *small*, *moderate*, and *large* extents revealed a statistically significant difference ($p < .01$) in Grade 4 assessment average scale scores (lower) in 2005 compared to no professional development. However, there were no other significant differences for the other data collection periods. Teacher professional development in use of computers or other technology in mathematics instruction revealed statistically significant differences ($p < .01$) in Grade 4 assessment average scale scores in 2013 for a *small* extent (higher scores) and 2005 for a *large* extent (lower scores). Teacher professional development in the effective use of calculators in mathematics instruction in *moderate* and *large* extents revealed the majority of significant differences ($p < .01$) in Grade 4 assessment average scale scores (lower scores) for a majority of years. Bennison and Goos (2010) identified a significant relationship between teachers' participation in professional development and both confidence in using the technology (computers and graphics calculators) and beliefs about the technology having a positive impact on students' math learning abilities. Goos and Bennison (2008) found the key concern of teachers regarding professional development related to technology integration was "how to integrate technology into classroom teaching in ways that improve students' understanding of mathematical concepts" (p. 126). The missing item is the technological pedagogical knowledge that teachers need to have to implement what they learned in the professional development. Researchers will provide some implication for research and practices in the next section.

7 Implications

The main ideas in successful development of educational technologies include: research concerning student thinking trajectories, proper representation of mathematical figures, technology experiences that carry over into off-tool studies, consideration of the implications of using a certain tool at a certain grade level, research supporting integration of tools into curricula, accurate mathematical expression, and engaging students with math in a way that cannot occur without the technology. The mutual relationship between research and technology integration requires constant attention to future needs and anticipation of the ways technology could improve mathematics studies (Heid and Blume 2008).

Observing a Grade 2 math class, Spencer (2014), found that the teacher is being guided by the expectations given by policy-makers rather than what is best for her students. Spencer reevaluated whether there is a need for more and higher-quality professional development or if the policies regarding education are at fault. He suggests professional development provides teachers with the capabilities to do their best, but in the classroom, there is something holding them back. He argued that it is not a matter of teachers falling into old habits, but a fear that if they change the way they are teaching, they will not fulfill what is expected of them as employees. With demanding lesson-planning and frequent testing required, little room is left for teachers to reinvent their methodology and better engage their students. Constructivism was remolded by policies into “scripted curriculum and programmed learning systems” (p. 75). While teachers wait for school policies to catch up to what students need, they will just have to “find a loophole... to sneak in good instruction” (p.75). Technology must be effectively integrated for it to be beneficial for both teachers and students. This integration should include the technology, curriculum, and professional development as described by Roschelle et al. (2010), with emphasis on the TPACK framework for designing professional development (Herring et al. 2016). School leaders could also play a role in supporting the integration of technology (Heintzelman 2017). Moreover, it could be that professional development could be used to implement the backward design in planning curriculum that meets school leaders’ goals as described by Guskey (2014). Additionally, Hill et al. (2013) recommended that scholars “should execute more rigorous comparisons of professional development designs at the initial stages of program development” (p. 476) to build a professional-knowledge based professional development.

The results of this present study are consistent with previous studies including both positive effects on mathematics learning outcomes (e.g. Cheung and Slavin 2013), and negative impacts based on extended levels of use (e.g. OECD 2015). We offer the following suggestions based on the results of our analysis:

- 1) Student reported restricted used of computers resulted in better mathematics performance. The analysis regarding students’ uses of computers revealed students that used computers *once or twice a month* in mathematics class had significantly higher values in mathematics achievement than students that *never or hardly ever* used computers. Students that used computers *once or twice a week or every day or almost every day* had significantly lower values in mathematics achievement (Table 2). Our findings suggest that it is rather beneficial to practice a more “selective” use of technology and that implementing more technology or increasing frequency of use does not always imply better performance.
- 2) Our results regarding students’ use of calculators reported by teachers showed that *restricted use* of a calculator during mathematics lessons revealed statistically significant ($p < .01$) differences (higher) in Grade 4 assessment scale scores in all six data collection periods when compared to calculators *not permitted*. Students’ reported use of calculators almost every day revealed statistically significant ($p < .01$) differences (lower) in Grade 4 assessment scale scores in all six data collection periods when compared to calculators *never or hardly ever* used (Table 3). *Always* using a calculator for math tests as reported by teachers and students yielded significantly ($p < .01$) different (lower) Grade 4 assessment scale

scores for all data collection periods (Table 4). It also worth noting that there may be a discrepancy in the measure of level of use as interpreted by teachers and/or students. However, our recommendation for selective use is based on teachers' reported use as these are the qualified individuals managing the learning environment for the students.

- 3) Professional development can be “successful” in terms of increases in technological pedagogical knowledge for the teacher. However, this does not specifically translate to effective implementation of this new knowledge, as there may be some obstacles which would need further investigation in future studies. In general, it appears that more teacher professional development regarding calculator use revealed significant ($p < .01$) differences (lower scores) in Grade 4 assessment scale scores for multiple data collection periods (Table 5). Teachers might have been unsuccessful in implementing their learning from the professional development effectively based on environmental or institutional barriers outside the scope of this study. As indicated in previous research, future research needs to further define educational technology tools can support teachers and positively influence students' mathematics learning (Utterberg et al. 2017). The present study results might suggest that although these teachers participated in professional development because they value the use of technology, they may need more guidance to integrate technology effectively in the classroom, which has been previously confirmed (e.g. O'Neal et al. 2017).

8 Conclusion

The purpose of this study was to determine the impact of teachers' use of technology and professional development on students' NAEP mathematics academic achievement. Technology use in mathematics education has demonstrated positive effects of mathematics achievement (Cheung and Slavin 2013; Li and Ma 2010) as well as negative impacts regarding academic performance in general (OECD 2015). Increased level of use of computers in mathematics instruction negatively impacted mathematics academic achievement compared to never using computers. Selective use of computers, once or twice a month, revealed significant increases in mathematics academic achievement compared to never using computers. Restricted use of calculators for mathematics lessons and tests revealed significant differences ($p < .01$) in terms of higher scores. While the NAEP participant teachers received professional development on the use of calculators for mathematics education, the data analysis in this study did not reveal improvements in students' test scores at the Grade 4 level.

Technology has the capacity to “support new pedagogies that focus on learners as active participants with tools for inquiry-based pedagogies and collaborative workspaces” (OECD 2015 p. 3). Technologies continually evolve, requiring the ongoing development of teaching knowledge described by Baumert and Kunter (2006), combined with technological pedagogical knowledge described by Mishra and Koehler (2006). Evidence, based on the results of this study and the existing literature, leads us to hypothesize that for effective professional development to have a positive impact on student academic achievement in mathematics, a supportive context for teachers' use of technologies is necessary (Spencer 2014).

Compliance with ethical standards

Conflict of interest The authors declare there is no conflict of interest.

References

- Bakker, M., van den Heuvel-Panhuizen, M., & Robitzsch, A. (2015). Effects of playing mathematics computer games on primary school students' multiplicative reasoning ability. *Contemporary Educational Psychology, 40*, 55–71.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle kompetenz von Lehrkräften [Teachers' professional competence]. *Zeitschrift für Erziehungswissenschaft, 9*, 469–520. <https://doi.org/10.1007/s11618-006-0165-2>.
- Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter et al. (eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers, mathematics teacher education 8* (pp. 25–48). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-5149-5_2.
- Bennison, A., & Goos, M. (2010). Learning to teach mathematics with technology: A survey of professional development needs, experiences and impacts. *Mathematics Education Research Journal, 22*(1), 31–56.
- Bouck, E. C., Joshi, G. S., & Johnson, L. (2013). Examining calculator use among students with and without disabilities educated with different mathematical curricula. *Educational Studies in Mathematics, 83*(3), 369–385.
- Briggs, S., & Crompton, H. (2016). Taking advantage of MOOCs in k-12 education: A blended approach. In D. Parsons (Ed.), *Mobile and blended learning innovations for improved learning outcomes* (pp. 297–309). Hershey: IGI Global. <https://doi.org/10.4018/978-1-5225-0359-0.ch015>.
- Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). *Effectiveness of reading and mathematics software products: Findings from two student cohorts*. Washington, DC: Institute of Education Sciences.
- Carr, J. M. (2012). Does math achievement h'app'en when ipads and game-based learning are incorporated into fifth-grade mathematics instruction? *Journal of Information Technology Education, 11*, 269–286.
- Cheung, A. C., & Slavin, R. E. (2013). Review: The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 988–1113. <https://doi.org/10.1016/j.edurev.2013.01.001>.
- Dash, S., de Kramer, R. M., O'Dwyer, L. M., Masters, J., & Russell, M. (2012). Impact of online professional development on teacher quality and student achievement in fifth grade mathematics. *Journal of Research on Technology in Education, 45*(1), 1–26.
- Dynarski, M., Agodini, R., Heavyside, S. N., Carey, N., Campuzano, L., Means, B., et al. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort*. Washington, DC: Institute of Education Sciences.
- Ernest, P. (1989). The knowledge, beliefs, and attitudes of the mathematics teacher: A model. *Journal of Education for Teaching, 15*, 13–33.
- Goos, M., & Bennison, A. (2008). Surveying the technology landscape: Teachers' use of technology in secondary mathematics classrooms. *Mathematics Education Research Journal, 20*(3), 102–130.
- Guskey, T. R. (2014). Planning professional learning. *Educational Leadership, 71*(8), 10.
- Handal, B., Campbell, C., Cavanagh, M., Petocz, P., & Kelly, N. (2013). Technological pedagogical content knowledge of secondary mathematics teachers. *Contemporary Issues in Technology and Teacher Education, 13*(1) Retrieved from <http://www.citejournal.org/vol13/iss1/mathematics/article1.cfm>.
- Heid, M. K., & Blume, G. W. (2008). Algebra and function development. In M. K. Heid & G. W. Blume (Eds.), *Research on technology and the teaching and learning of mathematics: Research syntheses* (Vol. 1, pp. 55–108). Charlotte: Information Age.
- Heintzelman, S. C. (2017). Integrating technology to engage students with EBD: A case study of school leader support. *Theses and Dissertations - Education Science, 28* Retrieved from https://uknowledge.uky.edu/edsc_etds/28.
- Herring, M. C., Koehler, M. J., & Mishra, P. (Eds.). (2016). *Handbook of technological pedagogical content knowledge (TPACK) for educators* (2nd edn.). New York: Routledge.

- Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher*, 42(9), 476–487.
- Horn, M. B., & Mackey, K. (2011). Transforming American education. *E-Learning and Digital Media*, 8(2), 133–144.
- Kennedy, M. (2010). Connecting to the future. *American School & University*, 82(9), 14–21.
- Kim, S., & Chang, M. (2010). Computer games for the math achievement of diverse students. *Educational Technology & Society*, 13(3), 224–232.
- Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education*, 9(1) Retrieved from <http://www.citejournal.org/vol9/iss1/general/article1.cfm>.
- Koehler, M. J., Mishra, P., & Yahya, K. (2007). Tracing the development of teacher knowledge in a design seminar: Integrating content, pedagogy and technology. *Computers & Education*, 49(3), 740–762.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820.
- Lee, Y.-H., Waxman, H., Wu, J.-Y., Michko, G., & Lin, G. (2013). Revisit the effect of teaching and learning with technology. *Educational Technology & Society*, 16(1), 133–146.
- Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, 22(3), 215–243. <https://doi.org/10.1007/s10648-010-9125-8>.
- Means, B. (2010). Technology and education change: Focus on student learning. *Journal of Research on Technology in Education*, 42(3), 285–307.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
- National Assessment Governing Board. (2007). *Assessment and item specifications for the NAEP 2009 mathematics assessment*. U.S. Department of Education. Washington, DC: U.S. Government Printing Office.
- National Assessment Governing Board. (2012). *Mathematics framework for the 2013 national assessment of educational progress*. U.S. Department of Education. Washington, DC: U.S. Government Printing Office.
- National Center for Education Statistics. (2006). *National assessment of educational progress: The nation's report card: Mathematics 2005 (NCES 2006–453)*. National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (2007). *Mathematics 2007: National assessment of educational progress at grades 4 and 8 (NCES 2007–494)*. National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (2010). *Mathematics 2009: National assessment of educational progress at grades 4 and 8 (NCES 2010–451)*. National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics (2011, March 25). NAEP Assessment Sample Design. Retrieved from https://nces.ed.gov/nationsreportcard/tdw/sample_design/
- National Center for Education Statistics. (2012). *Mathematics 2011: National assessment of educational progress at grades 4 and 8 (NCES 2012–458)*. National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (2013). *The nation's report card: A first look–2013 mathematics and reading (NCES 2013–456)*. National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics (2014). NAEP scoring. Retrieved from <https://nces.ed.gov/nationsreportcard/tdw/scoring/>
- National Center for Education Statistics. (2015a). *The condition of education 2015 (NCES 2015–144)*. National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics (2015b). About the 2015 mathematics assessment. Retrieved from https://www.nationsreportcard.gov/reading_math_2015/#mathematics/about
- National Center for Education Statistics (2015c). Questionnaires for students, teachers, and schools. Retrieved from <https://nces.ed.gov/nationsreportcard/bgquest.aspx>.
- National Center for Education Statistics. (2016a). *The condition of education 2016 (NCES 2016–144)*. National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.

- National Center for Education Statistics (2016b). NAEP data explorer [Data file]. Retrieved from <https://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx>
- National Center for Education Statistics (2017). *Frequently asked questions about NAEP survey questionnaires*. Retrieved from https://nces.ed.gov/nationsreportcard/bgquest_faq.aspx. Accessed 08 Feb 2017.
- O'Neal, L. J., Gibson, P., & Cotten, S. R. (2017). Elementary school teachers' beliefs about the role of technology in 21st-century teaching and learning. *Computers in the Schools*, 34(3), 192–206.
- Organisation for Economic Co-Operation and Development [OECD] (2015). *Students, computers and learning: Making the connection*. Programme for International Student Assessment, OECD Publishing. <https://doi.org/10.1787/9789264239555-en>.
- Polly, D. (2008). Modeling the influence of calculator use and teacher effects on first grade students' mathematics achievement. *Journal of Computers in Mathematics and Science Teaching*, 27(3), 245–263.
- Polly, D. (2011). Developing teachers' technological, pedagogical, and content knowledge (TPACK) through mathematics professional development. *International Journal for Technology in Mathematics Education*, 18(2), 83–95.
- Roberts, D. M. (1980). The impact of electronic calculators on educational performance. *Review of Educational Research*, 50(1), 71–98.
- Roschelle, R., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., & Gallagher, L. P. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal*, 47(4), 833–878.
- Shulman, L. S. (1986). Those who understand: A conception of teacher knowledge. *American Educator*, 10(1), 9–15.
- Spencer, J. T. (2014). PD isn't the problem. *Phi Delta Kappan*, 95(7), 74–75.
- Thompson, A. D., & Sproule, S. L. (2000). Deciding when to use calculators. *Mathematics Teaching in the Middle School*, 6(2), 126.
- Tienken, C. H., & Wilson, M. J. (2007). The impact of computer assisted instruction on seventh-grade students' mathematics achievement. *Planning and Changing*, 38(3), 181–190.
- Utterberg, M., Lundin, J., & Lindström, B. (2017). Conditions influencing mathematics teachers' uptake of digital tools – a systematic literature review. In P. Resta & S. Smith (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 2016–2029). Austin: Association for the Advancement of Computing in Education (AACE) Retrieved from <https://www.learnlib.org/p/177494/>.
- Van de Walle, J. A., Karp, K. S., & Williams, J. M. B. (2007). *Elementary and middle school mathematics. Teaching development*. Boston: Pearson.
- Walcott, C., & Stickles, P. R. (2012). Calculator use on NAEP: A look at fourth- and eighth-grade mathematics achievement. *School Science and Mathematics*, 112(4), 241–254.
- Yurdakul, I. K., Odabasi, H. F., Kilicer, K., Coklar, A. N., Birinci, G., & Kurt, A. A. (2012). The development, validity and reliability of TPACK-deep: A technological pedagogical content knowledge scale. *Computers & Education*, 58(3), 964–977.