TABLE OF CONTENTS

PREFACE

v *Brandon Milonovich, Teachers College, Columbia University*
Elizabeth Wentworth, Teachers College, Columbia University

ARTICLES

1 **Pokémon Battles as a Context for Mathematical Modeling**
William McGuffey, Teachers College, Columbia University

7 **A Reconceptualized Framework for 'Opportunity to Learn' in School Mathematics**
Temple A. Walkowiak, North Carolina State University;
Holly H. Pinter, Western Carolina University;
Robert Q. Berry, University of Virginia

19 **Comparing the Major Definitions of Mathematics Pedagogical Content Knowledge**
Jeffrey Johnson, Arizona State University

29 **About the Authors**

30 **Acknowledgement of Reviewers**

JOURNAL OF **MATHEMATICS EDUCATION** AT TEACHERS COLLEGE | SPRING 2017 | VOLUME 8, ISSUE 1

TEACHERS COLLEGE | COLUMBIA UNIVERSITY

Comparing the Major Definitions of Mathematics Pedagogical Content Knowledge

Jeffrey Johnson
Arizona State University

ABSTRACT The purpose of this research was to compare two definitions and assessments of Mathematics Pedagogical Content Knowledge (PCK). As part of a larger study, nine current and future teachers took an online version of the Measures of Knowledge for Teaching (MKT) - Mathematics assessment and the Cognitively Activating Instruction in Mathematics (COACTIV) assessment. Participants answered questions that demonstrated their understanding of students' misconceptions, ability to utilize multiple representations and methods, and knowledge of materials that they may be using for instruction. The data were analyzed, and the results indicate that teachers' scores on these two tests are correlated, but that the individual areas of knowledge were not.

KEYWORDS *pedagogical content knowledge, cognitively activating instruction, mathematics knowledge for teaching*

Introduction

Most researchers believe that pedagogical, content, and especially pedagogical content knowledge, or PCK, have positive effects on student success (Baumert et al., 2010; Campbell et al., 2014). The definition and description of Mathematical Pedagogical Content Knowledge has been made by several different groups with key centers in the United States (Ball & Forzani, 2011) and Germany (Baumert et al., 2010). Both of these groups combined definitions of Mathematics PCK with assessments to measure it among mathematics educators, and both tools have been tested to show a positive correlation between individual teachers' levels of PCK and student success on standardized tests.

These two descriptions of mathematics PCK are made by authors that held different theoretical perspectives to accommodate the interplay of internal and external change. Ball is generally considered a constructivist who believes that mathematics teachers develop knowledge for teaching through "pedagogical deliberations" (Ball, 1993). She claimed that teachers need a "bifocal perspective" to perceive the mathematics they are teaching about and the mind of the child they are teaching it to. In this framework, learning of PCK comes through reflection on what is to be taught and to whom it is to be taught. Baumert and the COACTIV group believe that teaching is a cognitive activity (Kunter et al., 2013) that is developed through specific training and is not significantly improved upon through the practical experience teachers have during their career. They emphasize their belief that teacher learning is not achieved through socialization into the profession, nor should it be studied based on the individual constructs of the knowledge teachers gain. In their framework learning occurs in formal teacher training when learners are taught things that they will later apply in their classrooms.

Both of their assessments have been shown to be valid and reliable, and their descriptions use similar language to describe the effect that this knowledge should have in the classroom. However, the assessments have not been compared to see how compatible they are with each other. There has not been a systematic effort to see if the different research centers in the field are discussing and measuring the same things. As part of a larger study on the development of Mathematics PCK, data were collected from a group of educators measuring their mathematics knowledge for teaching on these two different assessments. Their responses were examined to look for correlation between the assessments as a whole as well as individual sections of the two tests.

Background

Shulman (1987) provided a listing of seven categories of teacher knowledge that must be connected to practice. In later work, he would group them differently; however, these categories provide us with an avenue to separate areas of knowledge and the practices that employ and demonstrate them. Three of them, content knowledge, pedagogical knowledge, and pedagogical content knowledge, have been referred to as the core dimensions of teacher knowledge, and occupy a large amount of the literature in the field of mathematics education.

Content and, to a lesser extent, pedagogical knowledge have been the focal points on research into teacher knowledge and learning for decades. However, the studies of Begle (1979) and Monk (1994) found that while content courses and education courses taken by a mathematics teacher may have some effect on student learning, "it appears that courses in undergraduate mathematics pedagogy contribute more to pupil performance gains than do courses" in those other areas (Monk, 1994). One possible reason for this is the idea that while content courses and educational courses provide content and pedagogical knowledge—subjects that future teachers may have already been learning over years of observation—mathematics methods courses provide teachers with a more focused study of pedagogical content knowledge. Shulman (1987) defined pedagogical content knowledge as "the blending of content and pedagogy into an understanding of how particular topics, problems or issues are organized, represented, and adapted to the diverse interests and abilities of learners, and presented for instruction." (p. 8). This concept is also referred to as "Craft Knowledge" (Grimmett & Mackinnon, 1992) and has been applied to mathematics education using the phrase "Mathematical Knowledge for Teaching" (Ball et al., 2001).

Hill, Ball, and Schilling (2008) defined PCK as having three components; *Knowledge of Curriculum, Knowledge of Content and Students*, and *Knowledge of Content and Teaching.* They constructed this definition so that all of the areas were distinct from simple content knowledge, because "a teacher might have strong knowledge of the content itself but weak knowledge of how students learn the content or vice versa." Thus, in their construct Content Knowledge (CK) and PCK are separate aspects of teacher knowledge that can be developed independently. Hill et al. (2008) then constructed their assessment, the Measures of Knowledge for Teaching Mathematics or MKT, and attempted to measure teachers' knowledge in these areas. Originally the MKT was focused on Elementary School teachers and the knowledge they would use in teaching 4th and 5th grade mathematics classes. It was later expanded to include Middle School Material, and most recently into High School with the development of the Algebra 1 assessment.

Within each sub-section of the Algebra 1 MKT Assessment, there were multiple questions to assess teachers' knowledge; however, there may be overlap in the knowledge required. In the *Understanding Students* section, participants are asked to anticipate what might cause students to have difficulty understanding a problem and look at students' work to identify what caused their errors. For example, one question states that Mr. Anderson gave his student the problem

$$\frac{2(a+1)}{3a} + 3 - \frac{2}{3a} - \frac{6(a-2)}{3a} =$$

and one student showed the following (Figure 1):

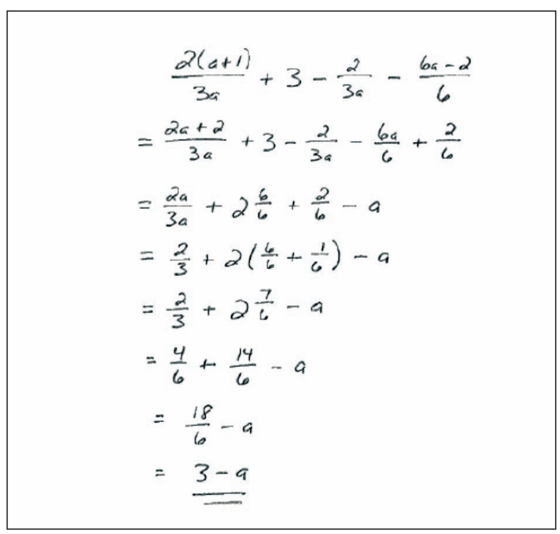

Figure 1. Question on Identifying Student Errors on MKT. Gates Foundation and ETS, 2012.

Participants are given four choices for what caused the student's error:

A) This student used the distributive property incorrectly

B) This student confounded mixed fractions with factors.

C) This student forgot to cancel common factors in several places.

D) This student needs to apply a more formal procedure by finding the common denominator and then adding all terms.

with the correct answer being B.

The *Utilizing Multiple Representations* section requires participants to look at problems and identify correct methods of solving them. One problem presents four different methods that students used to solve $-5x + 8 = 13x - 10$. Participants decided whether the work provided evidence that the student reasoned correctly (Figure 2).

	Provides Evidenceof Correct Student Reasoning	Does Not Provide Evidence of Correct Student Reasoning
A $-5x + 8 = 13x - 10$ $8 = 18x - 10$ $18 = 18x$ $1 = x$		
B $-5x + 8 - 13x + 10 = 13x - 10 - 13x + 10$ $-5x - 13x + 8 + 10 = 0$ $-18x + 18 - 18 = 0 - 18$ $\frac{1}{-18} \cdot \frac{-18x}{1} = \frac{-18}{1} \cdot \frac{1}{-18}$ $x = 1$		
C $-5x + 8 = 13x - 10$ $-5x + 8 = 3x$ $+5x \quad +5x$ $\frac{8}{8} = \frac{8x}{8}$ $1 = x$		
D $-5x + 8 = 13x - 10$ $-13x - 8 \quad -13x - 8$ $\frac{-18x}{-18} = \frac{-18}{-18}$ $x = 1$		

Figure 2. Question on Utilizing Multiple Representations on MKT. Gates Foundation and ETS, 2012.

In this problem (Figure 2), students A, B and D all showed acceptable work, but student C made an error. To answer this correctly, participants not only need to know three different methods for solving this problem, but also that you can only combine like terms. The recognition that a student incorrectly combined two terms might be a sign of understanding students just as much as it is a component of understanding methods.

A similar overlap may exist with the last section, *Understanding Curriculum and Instruction.* This section asks teachers to recognize different instructional materials and the benefits or challenges they may face while using them. For example, one question refers to a teacher using a geoboard to model slope and a description of a geoboard is included explaining that they are blocks with pins sticking out in a 1 inch grid pattern and are accompanied by rubber bands that can be stretched between the pins to create lines or polygons. In the problem, a student asks the question "Since the diagonal of one of the unit squares has length √2, does that mean you can make a line segment with slope √2 on the geoboard?" Four student responses are provided, and the participant is asked which statement gives the best insight into the question. The statements are:

Andy: Edward's right that the diagonal of the unit square has length √2, but its slope is 1.

Beth: Well, that doesn't matter. We can just turn the geoboard so that the diagonal is horizontal, and then we can see squares with side length √2.

Caitlin: Sure, but the square roots of two would just cancel. I think they always would, so you can't get √2 as a slope.

Dan: That's not right, because we can make one length of √2 and another length of 1 and use them as the rise and run. (Gates Foundation and ETS, 2012)

This question is designed to test teachers' familiarity with and understanding of geoboards, as evidenced by Caitlin giving the correct answer. It may be possible to answer this question by thinking through the methods described in the answers, reasoning through the limitations of shapes on a geoboard, and identifying which is the most insightful.

In contrast to Ball et al's. work (2001), Baumert's group (Krauss et al., 2008) felt that Content Knowledge was a necessary precondition to developing PCK. They utilized Shulman's 1986 original definition of PCK that it "includes knowledge on how best to represent and formulate the subject to make it comprehensible to others, as well as knowledge on students' subject-specific conceptions and misconceptions." (p. 9). The groups then added a third component of PCK based on research of effective mathematics instruction, namely the appropriate use of tasks as a means of developing students foundational understanding of the mathematics they are learning. Thus their definition of PCK also has three areas; *Knowledge of Mathematical Tasks for Learning, Understanding of Students Conceptions* and *Misconceptions, and Knowledge of Appropriate Mathematics-Specific Instructional Methods.* They also created an assessment, known as Professional Competence of Teachers Cognitively Activating Instruction and the Development of Student's Mathematical Literacy or COACTIV, and by including questions about both content and pedagogical content were able to show that their measurement of PCK was distinct from CK.

This assessment tries to draw a more distinct difference between the categories, even though the questions they are based on may be related. One section of the test starts with the statement "Many students have difficulty accepting the definition $a^0 = 1$." A question that tests *Understanding of Students* asks, "What might be the reasons for this? List as many as possible." The follow up question falls into the *Understanding Representations* section by asking participants to "outline as many ways (methods) as possible to make this definition accessible to students."

Another item begins with the statement, "There are S students and P professors at a university. There are six students to a professor." The problem states that the most common error that students made in representing this problem algebraically was writing "P=6S". A question assessing the *Understanding of Students* asks participants to "Please give possible reasons for this error being made—what might the students have been thinking?" This is followed by an *Understanding Tasks* question where teachers are asked to "Please briefly describe possible didactic interventions targeting this error." There are also additional questions related purely to content knowledge, such as "Please prove that $\sqrt{2}$ is irrational" and "Prove that the base angles of an isosceles triangle are congruent." In designing the COACTIV assessment, it was assumed that this type of knowledge was separate from PCK, and the questions were used to test for a correlation between content and pedagogical content knowledge.

These two groups from the United States and Germany are noteworthy because of their attempts to study their constructs of PCK through assessments and analysis of the results, and the continued use of those assessments by both of these groups and other researchers studying the topic. However, these two tests are very different. COACTIV is an open-ended assessment that was developed in Germany and tested on over 200 secondary mathematics teachers. It was demonstrated that higher levels of CK and PCK—as measured on the assessment—were correlated with increased student achievement on standardized assessments (Krauss et al., 2008). This research study also indicated that variations of teachers' knowledge were determined in their teacher preparation program and remained relatively fixed throughout the remainder of their teaching career.

In contrast, MKT is a multiple choice assessment that also correlates well with measures of teacher quality and student achievement (Hill, Umland, Litke, & Kapitula, 2012). Developed at the University of Michigan, it has been used for over ten years at sites across the U.S and has been split into elementary (K-5) and middle school (6-8) levels, with the recent introduction of a High School level (9) Algebra assessment. While the test shows overall validity, Hill et al. (2012) demonstrated that using cut scores of the quartiles allows for useful grouping of teachers. Those in the upper quartile have significantly higher quality lessons and student achievement than those in the lower quartile, while those in the middle two quartiles showed greater variation. Both of these tests have traditionally been administered using paper and pencil with a proctor observing the test takers. There is some evidence that online versions of assessments provide equivalent results (Weigold, Weigold, & Russell, 2013).

So while we have a good idea of the importance of PCK for Mathematics teachers, we do not know how the two main frameworks of Pedagogical Content Knowledge compare. If MKT as described by Ball et al. (2001) and PCK as defined by Baumert et al. (2010) are describing the same concept, then there should be a high level of correlation between the assessments that they have made. There should also be a correlation between individual Content Knowledge and their PCK, and the matching subsections.

Method

In January of 2016, students majoring in mathematics education at two traditional teacher preparation programs and a random selection of current teachers, including teachers at a large high school district and a small group of charter schools, were asked to participate in a research project. Students and teachers were sent an e-mail asking them to complete an online survey based on the questions from the MKT. The MKT assessment is accompanied by an answer key to allow for number-right scoring. However, to account for the possibility of participants guessing the answer to the questions, a formula scoring method was used with 3 points for a correct answer to a multiple choice question and 1 point for binary choice question, with incorrect answers scored as –1 points and skipped questions receiving 0. The formula scoring model has been debated in the past, with Lord (1975) claiming that it is based on an "indefensible assumption" that examinees either know or don't know an answer, and van Mameren and van der Vleuten (1999) claiming it shows increased bias. However they and others have found that use of a formula score may result in increased reliability of the assessment (Espinosa & Gardeazabal, 2010). This system also allows us to establish zero as the score a person with no pedagogical content knowledge would receive.

After the data from the MKT was collected, those who indicated that they were willing to answer additional questions were invited to participate in phase two of the study. Nine mathematics educators took a longer online survey that asked questions based on the COACTIV assessment. The COACTIV Assessment was also created with a Code Book to allow for standardized scoring. This system awards multiple points for questions that have multiple correct answers, while other questions may only be worth a maximum of 1 point. For consistency sake, this assessment was scored using the formula scoring model with 1 point for each correct answer, zero points for skipped questions, and – 1 for incorrect answers.

The purpose of this study was to determine to what extent the MKT and COACTIV assessments demonstrate concurrent validity. To accomplish this, I first ran a simple linear regression between the aggregate scores of those who took the COACTIV assessment and their scores on the MTK and calculated the Pierson product-moment correlation coefficient. While the scores on these two assessments are definitely ordered they are not necessarily continuous, which does bring into question the validity of using a test that requires interval data. However, the concept of knowledge is continuous, and other analysis of this type of data has shown it to be normally distributed. I also wanted to assess the degree to which the overlapping sub-constructs in the two instruments correlate, and the extent to which the non-overlapping sub-constructs in the instruments might improve the comprehensiveness and construct validity of each. To do this I took the individual scores on the three PCK sections from each assessment and the CK section of the COACTIV assessment and calculated the Pierson product-moment correlation coefficient.

Results

Nine participants completed the MKT and COACTIV assessments. Their demographic information is located in Table 1.

This group was relatively evenly split between mathematics (44%), mathematics education (33%), and non-education majors (22%). One pre-service and eight current teachers, ranging from having 2 to 25 years of experience, completed both assessments. 55% of participants came from University A, with the other four participants having attended four different programs. Other than the number of pre-service participants, the demographics for this group appears similar to that of the entire sample from the larger study of MKT scores. The MKT and COACTIV formula scores for these participants is located in Table 2.

Table 1
Participants

Participant	College Attended	Years of Experience	Major
1	University A	12	Mathematics
2	University A	12	Mathematics Education
3	University A	8	Mathematics Education
4	Online University A	13	Mathematics
5	Out of State University A	2	Geography
6	University A	–1 (Student Teacher)	Mathematics Education
7	Online University B	15	Criminal Justice
8	University A	12	Mathematics
9	Out of State University B	25	Mathematics

Table 2

MKT and COACTIV Scores

MKT	Strategy	Student	Curriculum	COACTIV	Content	Strategy	Student	Tasks
25	6	12	7	10	2	9	–3	2
31	22	10	–1	22	7	7	5	3
47	24	16	7	23	8	8	5	2
41	16	18	7	18	8	9	0	1
33	14	12	7	27	9	8	6	4
23	14	–6	15	5	0	5	0	0
33	14	16	3	9	0	3	3	3
35	22	10	3	4	1	2	2	–1
3	6	2	–5	–24	–10	–6	–7	–1

To test the relationship between the two tests, a *Pearson Correlation* test was run between the overall MKT and COACTIV test scores. While the data from the sample is not normally distributed, our analysis of the overall MKT scores showed a slight skew, which should not disqualify it from this test (Chok, 2010). The correlation coefficient was calculated as .840 with a two-tailed significance of .005, indicating that there is a positive correlation between the scores of the participants on the two tests with an r^2 of 0.70.

To investigate the possible relationships between the different sections of the test correlations were calculated for participants' scores for each section (see Table 3). Based on these results the only subsections that are correlated are the *Understanding Strategy* section of the MKT and the *Understanding Students* section of the COACTIV ($r = .771$, $p = .015$). None of the corresponding sections were significantly correlated. *Understanding Students* from the two tests had an r of .426 with p of .253, and Understanding Strategy had an r of .338 with p of .373.

Additionally, the Content Knowledge questions on the COACTIV assessment were strongly correlated with overall MKT scores. This does seem to validate the claim by Baumert and others that Content Knowledge and Pedagogical Content Knowledge, while separate areas, are highly correlated (Klickmann et al., 2015).

Table 3

Correlations Between Scores on Sections of MKT and COACTIV

	MKT	Strategy	Student	Curriculum	COACTIV	Content	Strategy	Student	Tasks
MKT	1								
Strategy	.752	1							
Student	.697	.310	1						
Curriculum	.403	.089	–.161	1					
COACTIV	.840*	.572	.559	.442	1				
Content	.856**	.600	.580	.419	.979**	1			
Strategy	.745*	.338	.474	.616	.908**	.902**	1		
Student	.776*	.771*	.426	.251	.856**	.801**	.584	1	
Tasks	.438	.133	.526	.108	.762*	.660	.627	.670*	1

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

Conclusions and Implications

Even though Baumert et al. (2010) and Ball et al. (2001) used different words to describe their components of Mathematical Pedagogical Content Knowledge, Depaepe, Verschaffel, and Kelchtermans (2013) identified two areas, *Understanding of Students* and *Understanding of Strategies,* which were conceptually identical for both. Likewise, the third areas, *Knowledge of Curriculum* and *Knowledge of Tasks,* seemed to overlap, seeing as curriculum can be broken down into tasks put into a specific order. According to both authors, these descriptions of PCK came before the development of their assessments; thus, the test should be a good measure of the definitions. Given that the scores of the participants on the two tests are correlated with a coefficient of .840 (p = 0.005), we can conclude that the tests are measuring mostly the same thing. In effect, 70% of the variation of scores on one is accounted for by the variation of scores on the other, meaning that 30% of the variation is different. Since they both claim to be testing for Mathematical Pedagogical Content Knowledge, we can conclude that the two definitions of PCK are mostly the same.

An ideal explanation of the variance would be that while two sections of each are described using similar language, the third area in each is different. However, none of the scores for the individual components are significantly correlated with their counterparts. This may be due to lack of participants, the length of the COACTIV assessment which lead to a higher number of unanswered questions, overlap in the area of knowledge being tested by each question, or because the type of questions that were used to assess the separate type of knowledge were slightly different. Additionally, this difference could be caused by the differences in the initial construction of the test. The MKT and COACTIV were created in two different languages and were tested with teachers working with very different types of students. In the *Understanding of Students* section, for example, issues like working with second language learners may be highly critical for one group but less important for the other.

However, knowing that these two groups created tests that measure the same type of knowledge should mean that the results based on one are similar to the results based on the other, and can help to strengthen the claims made about the value of PCK. Both groups had previously shown a relationship between a teacher's level of PCK—as measured on their assessment—and their students' achievement. Knowing that the assessments are similar provides greater evidence of that relationship. Additionally, Baumert's group had found that Content Knowledge was a separate but correlated area of knowledge (2010). These results support this claim, as CK from the COACTIV assessment was highly correlated with scores on the MKT (.856).

Another comparison can be made between how the creators of these assessments believe this knowledge is gained. Both Baumert et al. (2010) and Ball et al. (2001) believe that Content Knowledge of Mathematics is a prerequisite of PCK, and may be a subset of it. Thus, it would make sense that teachers who had majored in mathematics would have higher levels of PCK than those from other fields because of the increased number of mathematics courses they had taken in college. However the groups with the highest levels of PCK were mathematics education majors (Mean COACTIV score of 16.7, MKT of 33.7) and those from other fields (Mean COACTIVE score of 18, MKT of 33), both having average scores higher than those who majored in mathematics (Mean COACTIVE score of 2, MKT of 26). To get a B.S. in mathematics at University A, a student begins by taking Calculus I and completes 14 total upper division mathematics classes to get their degree. However, most teachers at the secondary level don't teach Calculus or more advanced courses. Thus, while those classes are informative, they may not be useful in developing CK among secondary teachers.

Questions for Further Study

Knowing that the main assessments of mathematics pedagogical content knowledge are measuring the same things allows us to compare results made by both groups. However, there are many gaps in our knowledge that still need to be filled. For more effective teacher preparation, it would be useful to discover what sections within the framework of Pedagogical Content Knowledge are most valuable for teachers. We have evidence that pre-service teachers demonstrate the most growth in *Understanding of Students* and may improve in the other areas, but we do not know what classes or experiences are causing that growth, nor the value of that knowledge. While many of the studies of teacher knowledge and student achievement have used single scores for comparison, there may be specific components of PCK that have a greater effect. Knowing that would allow teacher educators to focus on those topics.

We also do not know how a teachers' knowledge displays itself in the classroom. There is evidence that teachers with higher levels of PCK teach better lessons according to the Measures of Quality Instruction, which were graded by researchers looking at video tapes of specific lessons (Hill, Umland, Litke, & Kapitula, 2012). However, they could not tell if a teacher's level of mathematical knowledge was noticeable by students, administrators or other observers. It would be useful to have secondary assessments of teacher's knowledge and behaviors from outside sources to use as a comparison between what they say about teaching mathematics and the manner in which they actually teach mathematics.

Moreover, we do not really know how big of a role teacher knowledge plays in student learning. It has been estimated that teachers account for between 1% and 14% of the variability in student improvement (American Statistical Association, 2014). If this is true, how much of the variability is described by the teachers' knowledge versus their beliefs, behaviors, or other demographic information? Most of the studies linking PCK to student achievement have been limited to end of course test results or assessments designed specifically to find the relationship, while studies of the value added by specific teachers rely on longitudinal data related to student achievement on standardized tests. It would be useful to link those two methods together to tease out the value of teacher knowledge on student success.

References

American Statistical Association. (2014). Statement on Using Value-Added Models for Educational Assessment, Executive Summary.

Ball, D., & Forzani, F. (2011). "Building a Common Core for Learning to Teach: And Connecting Professional Learning to Practice." *American Educator 35*(2) 17.

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. *Handbook of research on teaching, 4,* 433–456.

Ball, D. L. (1993). "Halves, pieces, and twoths: Constructing and using representational contexts in teaching fractions." *Rational numbers: An integration of research,* 157–195.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180.

Begle, E. G. (1979). *Critical Variables in Mathematics Education: Findings from a Survey of the Empirical Literature.* Mathematical Association of America

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Educational researcher,* 5–13.

Campbell, P. F., Nishio, M., Smith, T. M., Clark, L. M., Conant, D. L., Rust, A. H., & Choi, Y. (2014). The Relationship Between Teachers' Mathematical Content and Pedagogical Knowledge, Teachers' Perceptions, and Student Achievement. *Journal for Research in Mathematics Education, 45*(4), 419–459.

Carpenter, T. P., & Fennema, E. (1992). Cognitively guided instruction: Building on the knowledge of students and teachers. *International Journal of Educational Research, 17*(5), 457–470.

Chetty, R., Friedman, J., & Rockoff, J. (2014). Discussion of the American Statistical Association's Statement (2014) on Using Value-Added Models for Educational Assessment. *Statistics and Public Policy, 1*(1), 111–113.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood* (No. w17699). National Bureau of Economic Research.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2013). *Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood* (No. w19424). National Bureau of Economic Research.

Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education, 34,* 12–25.

Dash, S., Magidin de Kramer, R., O'Dwyer, L. M., Masters, J., & Russell, M. (2012). Impact of Online Professional Development or Teacher Quality and Student Achievement in Fifth Grade Mathematics. *Journal of research on technology in education, 45*(1), 1–26.

Espinosa, M. P., & Gardeazabal, J. (2010). Optimal correction for guessing in multiple-choice tests. *Journal of Mathematical Psychology, 54*(5), 415–425.

Gates Foundation and ETS. (2012). Measures of Effective Teaching Longitudinal Database, Teacher Knowledge Assessments. Downloaded from http://www.icpsr.umich.edu/cgi-bin/file?comp=none&study=34345&ds=8&file_id=1168527&path=METLDB

Grimmett, P. P., & Mackinnon, A. M. (1992). Craft knowledge and the education of teachers. *Review of research in education,* 385–456.

Hill, H. C., Umland, K., Litke, E., & Kapitula, L. R. (2012). Teacher quality and quality teaching: Examining the relationship of a teacher assessment to practice. *American Journal of Education, 118*(4), 489–519.

Hill, H. C., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for research in mathematics education*, 330–351.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for research in mathematics education,* 372–400.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American educational research journal, 42*(2), 371–406.

Hill, H. C., Umland, K., Litke, E., & Kapitula, L. R. (2012). Teacher quality and quality teaching: Examining the relationship of a teacher assessment to practice. *American Journal of Education, 118*(4), 489–519.

Kagan, D. M. (1992). Implication of research on teacher belief. *Educational Psychologist, 27*(1), 65–90.

Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S. Cheo, M., & Baumert, J. (2015). Content knowledge and pedagogical content knowledge in Taiwanese and German mathematics teachers. *Teaching and Teacher Education, 46,* 115–126.

Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology, 100*(3), 716.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Eds.). (2013). *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project.* Springer Science & Business Media.

Lord, F. M. (1975). Formula Scoring and Number-Right Scoring. *Journal of Educational Measurement, 12*(1), 7–11.

Marshall, J. H., & Sorto, M. A. (2012). The effects of teacher mathematics knowledge and pedagogy on student achievement in rural Guatemala. *International Review of Education, 58*(2), 173–197.

Metzler, J., & Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation. *Journal of Development Economics, 99*(2), 486–496.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of education review, 13*(2), 125–145.

Ng, W., Nicholas, H., & Williams, A. (2010). School experience influences on pre-service teachers' evolving beliefs about effective teaching, *Teaching and Teacher Education,* 26.

Ottmar, E. R., Rimm-Kaufman, S. E., Larsen, R. A., & Berry, R. Q. (2015). Mathematical Knowledge for Teaching, Standards-Based Mathematics Teaching Practices, and Student Achievement in the Context of the Responsive Classroom Approach. *American Educational Research Journal*, 787–821

Sadeghi, B., & Zanjani, M. S. (2014). The Role of Experience to Bring Association between Teachers' Professional Knowledge and Teaching Performance. *Journal of Foreign Languages, 2*(1), 177–200.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review, 57*(1), 1–23.

Silverman, J., & Thompson, P. W. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of mathematics teacher education, 11*(6), 499–511.

Speer, N. M., King, K. D., & Howell, H. (2015). Definitions of mathematical knowledge for teaching: using these constructs in research on secondary and college mathematics teachers. *Journal of Mathematics Teacher Education, 18*(2), 105–122.

Szydlik, J. E., Szydlik, S. D., & Benson, S. R. (2003). Exploring changes in pre-service elementary teachers' mathematical beliefs. *Journal of Mathematics Teacher Education, 6*(3), 253–279.

Tchoshanov, M. A. (2011). Relationship between teacher knowledge of concepts and connections, teaching practice, and student achievement in middle grades mathematics. *Educational Studies in Mathematics, 76*(2), 141–164.

Tillema, H. H. (2000). Belief change towards self-directed learning in student teachers: Immersion in practice or reflection on action. *Teaching and Teacher Education, 16*(5), 575–591.

Walton, E., Nel, N. M., Muller, H., & Lebeloane, O. (2014). 'You can train us until we are blue in our faces, we are still going to struggle': Teacher professional learning in a full-service school. *Education as Change, 18*(2), 319–333.

Weigold, A., Weigold, I. K., & Russell, E. J. (2013). Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods. *Psychological methods, 18*(1), 53.

Williams, M., & Burden, R. L. (2000). *Psychology for Language teachers: A Social Constructivist Approach.* Cambridge: Cambridge University Press.

van Mameren, H., & van de Vleuten, C. P. M. (1999) "The effect of a 'don't know'option on test scores: Number-right and formula scoring compared." *Medical Education 33*(4), 267–275.

Zheng, H. (2009). A review of research on EFL pre-service teachers' beliefs and practices. *Journal of Cambridge Studies, 4*(1)