Educ Stud Math (2017) 94:329–345
DOI 10.1007/s10649-016-9735-2

Teachers' pedagogical content knowledge and mathematics achievement of students in Peru

Santiago Cueto[1] · Juan León[1] · M. Alejandra Sorto[2] · Alejandra Miranda[1]

Published online: 20 October 2016

Abstract After improving enrolment rates significantly, many developing countries such as Peru are facing the challenge to increase learning levels among students. Over the past few years, many researchers have turned to teacher-related variables as a way to better understand classroom processes that may help increase learning levels among students. In this study, we analyze one of these, that falls under what Shulman (*Educational Researcher, 15* (2), 4–14, 1986) called pedagogical content knowledge (PCK). Specifically, in this study, we analyze one of the areas of PCK which is knowledge of content and students. This was measured through a test where teachers were asked to explain students' mistakes and predict responses in similar mathematics exercises. We explore if PCK is associated with the socioeconomic status of children and if it has an effect on children's achievement. Additionally, we analyze which teacher characteristics are associated with higher scores in PCK. The analysis uses the Young Lives longitudinal survey for Peru. We found that students' socioeconomic status at age 1 and maternal education were positively associated with their teachers' PCK by the time students were enrolled in fourth grade, thus depicting a very unequal education system. Furthermore, teachers' PCK was positively associated with student achievement, but only when a threshold for the PCK test was established. For our sample, male teachers, who were younger, and Spanish speakers had higher PCK scores.

✉ Santiago Cueto
scueto@grade.org.pe

Juan León
jleon@grade.org.pe

M. Alejandra Sorto
sorto@txstate.edu

Alejandra Miranda
amiranda@grade.org.pe

1 GRADE, Av. Grau 915, Barranco, Lima 4, Peru

2 Texas State University, San Marcos, TX, USA

Keywords Pedagogical content knowledge · Mathematics achievement · Peru · Socioeconomic status · Longitudinal study

1 Introduction

After achieving high rates of enrolment, many developing countries are facing the challenge of increasing learning levels among their students. For example, in Peru in 2012, the net enrolment rates for preschool, primary, and secondary education were 75, 93, and 81 %, respectively[1], with increasing tendencies at all levels over the past few years (higher education enrolment is also increasing). However, the national evaluation of second graders in 2012 showed that only around 13 % of students achieved at the satisfactory level in mathematics; the tendency over the past few years has been of very modest increases (Cueto, 2013). In secondary education, achievement is at similar low levels. For example, the 2012 Program for International Student Assessment (PISA) evaluation for students age 15 showed that in mathematics, almost 75 % of students scored in the bottom two levels of achievement (1 and below), and only around 9 % scored in levels 3 to 6 combined. This resulted in Peru having the lowest average among countries and regions participating in PISA[2]. Furthermore, both reports plus other evaluations have shown that socioeconomic status is a strong predictor of educational results (OECD, 2010, 2013; UMC, 2004, 2006). This makes Peru, like many other developing countries, highly unequal with regard to students' achievement.

These results raise a question for educational researchers: What variables from within the educational system could have an impact on both increasing averages and reducing inequalities associated with socioeconomic status? Recent studies in Peru and elsewhere have turned to teacher-related variables, with the hypothesis that teacher abilities and pedagogical practices in the classroom may be key to increasing learning (Creemers, 1994; Scheerens, 1990). This would include teachers' subject matter knowledge but also other types of knowledge (Kunter et al., 2013). Related to this, in this paper, we present data of one aspect of what has been described in the literature as Pedagogical Content Knowledge (PCK). PCK is an emerging area of research given that it is considered a critical condition of teachers' instructional quality (Baumert et al., 2010; Ngo, 2012). In developing countries, including Peru and Latin America, there is a lack of studies on this topic. In this paper, we attempt to fill this gap by exploring the association between teachers' PCK and student achievement. For the study, we use a longitudinal design, which is uncommon in developing countries. This design allows us to correlate socioeconomic indicators from when children were aged 1 year with the characteristics of their teachers several years later. It also allows us to estimate the association of PCK with students' achievement and to relate teacher characteristics with higher scores in PCK. In the following section, we briefly review the international literature related to PCK.

[1] Data estimated by authors based on data available from the Ministry of Education (http://escale.minedu.gob.pe/indicadores2011) on January 9, 2014.
[2] Data obtained from Peru's report on PISA, retrieved from http://www2.minedu.gob.pe/umc/PISA/Pisa2012/Informes_de_resultados/Informe_PISA_2012_Peru.pdf on January 9, 2014.

2 Literature review

2.1 Untangling the construct of pedagogical content knowledge

The construct of teachers' knowledge has been used in several studies since Shulman's paper (1986). There is some evidence, particularly from developing countries, that teachers' knowledge of subject matter they teach has an impact on student learning. For example, for Peru, Metzler and Woessman (2012) found that an increase of one standard deviation in teachers' knowledge of subject matter increased student achievement (reading and mathematics) by about 0.10 standard deviations. This study had no data on the pedagogical skills of teachers, but it does suggest that teachers' knowledge of subject matter by itself, at least in developing contexts such as the Peruvian, has an effect on student achievement. In industrialized countries such as the USA, there has been a discussion around whether teachers should master content at the level expected to be learned by their students or higher (e.g., as required to approve mathematics courses at the university level) in order to have the minimum knowledge necessary to detect student mistakes and misconceptions about mathematics (Conference Board of the Mathematical Sciences, 2001). However, in developing countries, even though there has been an expansion of teacher certification and training in the last decades, especially in Africa (Passos, Nahara, Magaia, & Lauchande, 2005), there is still evidence that teachers do not completely master the contents students are expected to learn; in other words, when administered items based on the curriculum, they often provide wrong answers. Studies developed in Brazil by Harbison and Hanushek (1992) and in Peru by Metzler and Woessman (2012) have shown this pattern.

However, even if teachers' knowledge of subject matter is relevant for learning, research suggests that other types of teacher knowledge also may contribute to promote student learning. In the literature, there is an understanding that domain-specific skills and also general pedagogical knowledge may have an effect on students' learning gains (Grossman & Schoenfeld, 2005; Hiebert, Morris, Berk, & Jansen, 2007). In 1986, Shulman suggested that not including diverse aspects of teachers' knowledge of subject matter they taught was a serious omission from previous conceptualizations and empirical studies aiming to understand student learning, with multiple implications for the development of theory and practice. He suggested that this knowledge could be further divided into subject matter content knowledge ("amount and organization of knowledge per se in the mind of the teacher," p. 9), PCK ("subject matter knowledge for teaching," p. 9), and curricular knowledge ("knowledge about programs designed for the teaching of particular subjects and topics at a given level," p. 10). Shulman suggested that knowledge of subject matter was a necessary but not sufficient requisite for high-quality teaching; teachers also needed knowledge of how to teach any given subject matter. This knowledge needed to take into account not only the curriculum and materials available but also the cognitive processes of students in specific contexts.

Depaepe, Verschaffel, and Kelchtermans (2013) made a systematic review of the way PCK is conceptualized in mathematics educational research and found that even though most authors refer to Shulman (1986) when they define PCK, there is no consensus in the literature about the components covered by it. Furthermore, Hill, Ball, and Schilling (2008) have argued that there is still lack of conceptual clarity on the constructs suggested by Shulman (1986). These authors have suggested a domain map for what they call "mathematical knowledge for teaching," further subdivided into knowledge of subject matter and PCK, each with its own subcategories. For these authors, PCK included knowledge of content and teaching and *knowledge of content and students*. The latter is defined as "content knowledge intertwined with knowledge of how students think about, know, or learn this particular content" (p. 375). This is more precisely what has been measured in the instruments administered to teachers included in this study, specifically in the field of mathematics.

Defining a construct conceptually is only the first step to measuring it. Several authors have attempted to measure diverse aspects of PCK, using a variety of methods. For example, Hill and Ball (2004) developed multiple-choice items aimed at measuring "teachers' common and specialized knowledge of content" (p. 337) in elementary number concepts and operations, etc. Their scales, administered to teachers, resulted in reliabilities above 0.70. The sample items provided by the authors presented teachers with classroom situations where they had to observe students' responses or reasoning and judge their value (although in the end, there was always only one correct response). Hill, Ball, and Schilling (2008) used multiple-choice items to measure PCK in mathematics and did a reanalysis of the data from the previous study (Hill & Ball, 2004) by using the same bank of items and discovered that often times, they loaded in more than one factor and the reliabilities of the scales were low. They suggested that a combination of multiple-choice and open-ended items could be the best way to capture teachers' knowledge of content and students. An, Kulm, and Wu (2004) used surveys and classroom observations to measure PCK among Chinese and US teachers.

Koirla, Davis, and Johnson (2008) have developed a performance assessment task, where prospective teachers are provided with pedagogical situations involving students' reasoning and responses and later used a rubric to score the answers. This type of instrument was devised to serve potentially both for measurement and as a training tool for prospective teachers. Lim-Teo, Chua, Cheang, and Yeo (2007) developed open-ended items to measure PCK in mathematics and found acceptable levels of inter-rater reliability. Park and Oliver (2007) have used qualitative methods to understand teachers' conceptualizations of PCK. Turnuklu and Yesildere (2007) developed a procedure, combining quantitative and qualitative methods, to test the abilities of pre-service primary teachers in Turkey. Even though they all claim to measure PCK, they often concentrate on one or a few specific aspects of the construct.

2.2 Unequal distribution of pedagogical content knowledge

We have been able to find a few studies exploring variables associated with teachers' PCK in developing countries. For example, Guadalupe, Leon, and Cueto (2013) explored the associated factors (at teacher and school) with teacher knowledge in math and reading comprehension with a sample of 697 sixth grade teachers in Peru. At the teacher level, the study found that younger, male, experienced, and regularly trained teachers had a higher knowledge than their peers; at the school level, teacher knowledge was positively associated with the average socioeconomic status of students and the type of school (full grade over multigrade).

In terms of unequal distribution of PCK, Borman and Kimball (2005) in Nevada, Guadalupe et al. (2013) in Peru and Ngo (2012) in Cambodia found that there was a positive association between teacher quality[3] and school socioeconomic status; in other words, schools that enroll students from high socioeconomic status had teachers with higher quality.

Although, teachers' PCK and knowledge in technology (TPACK) are not the same since the latter is a more specific knowledge; internationally, Koh, Chai, and Tsai (2014) explored what factors were associated with teachers' PCK and TPACK with a sample of 450 teachers (primary, secondary, and junior college) from Singapore. A correlational analysis performed by the authors indicated that younger, less experienced, male, and secondary/junior college school teachers had a higher TPACK than older, more experienced, female, and primary school teachers. However, only teaching level and teaching experience remained statistically significant once they held constant the other variables.

[3] They used a measure of pedagogical content knowledge or knowledge of a subject matter.

2.3 The relationship between pedagogical content knowledge and student achievement

There is some research on whether students of teachers with higher PCK have higher scores in mathematics achievement. Previous analysis suggests that there is a positive association between PCK and students' achievement, even when controlled for SES. For example, Hill, Rowan, and Ball (2005) found that mathematical knowledge for teaching was associated with achievement gains over a year for first and third graders in the USA, even when controlled for SES. Similarly, Rockoff, Jacob, Kane, and Staiger (2008) found that teachers' mathematics knowledge for teaching was associated with small but significant increases in scores for middle school students in the USA. Baumert et al. (2010) found that a measure of PCK is a significant predictor of student mathematics learning gains in German secondary schools; furthermore, they found that empirically they could differentiate teachers' knowledge of subject matter from PCK, with the latter having a stronger weight in predicting achievement. It is also interesting that through the use of structural equation models, these authors identified cognitive activation and curricular and cognitive level of task as mediating variables for PCK. Kunter et al. (2013) explored how professional competence (measured through different aspects, such as teachers' PCK, professional beliefs, work-related motivation, and self-regulation) impacts instruction and, in turn, student outcomes in mathematics in Germany. Marshall et al. (2009) found an association between teachers' PCK and mathematics achievement in Cambodian primary schools; interestingly, in this paper, the authors divided PCK skills of teachers in three levels, and the association with student achievement was highest for teachers in the highest PCK category. Ngo (2012) also found a positive effect of mathematics teachers' PCK on third-grade student achievement in Cambodia. PCK was measured through a questionnaire about teachers' knowledge of student tasks, knowledge of student misconceptions, and instructional practices. Marshall and Sorto (2012) used several aspects of mathematics teacher knowledge and found that it was associated with student achievement in Guatemala. Finally, Carnoy and Arends (2012) found a positive association between mathematical teachers' knowledge (which included teachers' knowledge of subject matter and PCK) and student achievement in Botswana.

In sum, most of the literature related to PCK and its effects on student achievement shows that it may be relevant. However, there is a scarcity of research on the effects of PCK on student achievement in Latin America. The existing studies used teacher knowledge of subject matter as a PCK measure (Marshall & Sorto, 2012; Metzler & Woessman, 2012). Finally, while some of the studies mentioned above used a repeated measures design (e.g., Hill et al., 2005), to our knowledge, this is the first longitudinal study with background data of children since they were aged 1 year and includes a control for preschool abilities and changes in socioeconomic variables over time.

3 Research questions

In this paper, we first explore whether knowledge of content and students, which as we said before is the component of PCK that we are measuring, is associated with early indicators of the socioeconomic status of children. In an unequal educational system, PCK would be positively associated with SES; in an equal opportunity, educational system the association would be zero, and in a system promoting a reduction of inequality, the association should be negative.

Second, we explore if there is a positive association between PCK and students' achievement (controlling for individual, family, teacher, and school-level variables). Finally, we explore which teacher characteristics are associated with higher scores in our PCK measure.

4 Methods

4.1 Design and data

The data used in this paper comes from the longitudinal study Young Lives. This is an international study that follows the lives of 12,000 children over 15 years in Ethiopia, India, Vietnam, and Peru. We used data from the three rounds of household surveys of Young Lives (administered in 2002, 2006, and 2009) along with information from the School Survey (administered in 2011) from Peru. While Young Lives has information on two cohorts (i.e., born in 1994 and in 2000), in this study, we only use data for a subsample of the younger cohort. The original sample in Peru was chosen randomly from 20 sites across the country, although the 5 % richest districts were excluded from the framework (Escobal & Flores, 2008).

Young Lives provides a unique opportunity to analyze the association between PCK and mathematics achievement, given that it provides measures of socioeconomic status at age 1 and children's skills before entering primary school and PCK and achievement measures of students by age 11 (i.e., fourth grade). Our measure of PCK was administered toward the end of the 2011 school year.

4.2 Sample

The younger cohort from Young Lives originally had 2052 children (in round 1), 1963 children in round 2, and 1943 in round 3 (Cueto, Escobal, Penny, & Ames, 2011). A subsample of 572 students in 132 primary schools was selected randomly and visited at school to inquire about their educational opportunities (Guerrero et al., 2013). In this study, we focus only on Young Lives children who were in fourth grade at the time of data collection and their mathematics teachers. Fourth grade was selected because it was the predominant grade in the sample. Table 1 includes information on the analytical sample, specifically on the types of schools children attended; private schools charge fees, while public schools are free. EIB schools are those with an intercultural bilingual (Indigenous-Spanish) education program. There are 312 students grouped in 102 schools; on average, there are two students per teacher.

4.3 Variables

Dependent variable Achievement of children was measured through a test that focused on number and number sense[4], given that previous studies show that these are topics covered widely in Peruvian classrooms (Cueto, Ramirez, & Leon, 2006). The test had 37 items, which were taken from national standardized tests (i.e., Census Evaluation of Peru in 2007 and National Evaluation in 2004) and the internal consistency was 0.91. In order to calculate and calibrate the achievement of students, the Young Lives study used Rasch analysis. The mean score was fixed at 300 and the standard deviation at 50.

Main independent variable To measure teachers' knowledge of content and students, we used a relatively simple and short instrument that was part of the School Survey. Specifically, we used a test formed by multiple-choice items where the teachers had to analyze a student's response. The items were chosen and adapted from internationally used instruments (Carnoy & Arends, 2012; Marshall &

[4] According to the National Council of Teachers of Mathematics (1989), number sense refers to the ability to understand the meaning of numbers, recognize the relative size of numbers, and use referents for measuring objects and events.

Table 1 Number of students, schools, and teachers by type of school

Type of school	Students	Teachers	Schools
Urban			
Private	16	10	10
Public	205	96	54
Rural			
Public Spanish	49	22	21
Public EIB	42	17	17
Total	312	145	102

Source: School Survey—Young Lives

Sorto, 2012; Sorto, Marshall, Luschei, & Carnoy, 2009). This test had 14 multiple-choice items[5] with typical fourth-grade students' mistakes in mathematics. All items cover common student errors in addition, subtraction, multiplication, and division of whole numbers and fractions, which cover a subset of the contents covered in the standardized test with students[6]. The answers were coded 1 if correct and 0 if blank or incorrect. We dropped four items with poor item-total correlation (below 0.10), for a final set of 10 items[7]. The internal consistency (Cronbach's alpha) for this reduced scale was 0.62; Hair, Anderson, Tatham, and Black (1998) suggest a minimum of 0.60 for internal consistency indexes. However, we realize that higher levels of internal consistency would be desirable, but as mentioned before, it is not uncommon among studies measuring PCK to find low indicators of reliability. We discuss later some potential actions for improving reliability in future instruments.

Student variables We controlled for socioeconomic status in 2002 (round 1 of Young Lives) by using the wealth index. This is a continuous variable formed by the aggregation of housing quality, consumer durables, and housing services available, which takes into account several dimensions associated to wealth. The changes in this indicator by 2009 (round 3 of Young Lives) were also included in the analysis (i.e., changes in the distribution among terciles of wealth). Maternal education was also included in the analysis, taking the value 1 if the mother had secondary education (incomplete or complete, about 55 % of the sample) or more and 0 if less than incomplete secondary education. Additionally, we controlled for the following individual variables: gender (1 if female), age in months, and mother tongue (1 if native language, 0 if Spanish). With regard to preschool abilities, we used the Child Development Assessment (CDA), administered by age 5 (round 2 of Young Lives). This is a test taken from an international study used by the International Evaluation Association, which measures mathematical notions such as few, more, none, and so on. It includes 15 items, and the score is a Rasch score with a mean of 300 and a standard deviation of 50, with a person reliability index of 0.56[8]. Mandatory age of enrolment in primary schools in Peru is 6 years.

School variables We also included school variables in the analysis, including if the school was public (1) or private (0) and school place of residence (1 if rural).

[5] The Appendix contains a sample item of the final set of 10.
[6] The PCK test does not include items on routine and real-life problems since they measure the application of math algorithms, and the focus of the PCK test was on how students solve math algorithms.
[7] The content of the test did not change after the elimination of the four items.
[8] For more information on this test and its psychometric characteristics, see Cueto et al. (2009)

Teacher variables Included gender (1 if female), age, mother tongue (1 if native), years of experience, and hours of in-service teacher training in the last 2 years (value of 0 for no training, 1 for 100 h or less, 2 for more than 100 and less than 200, and 3 for more than 200 h of training). The variable "teacher went to university" is a dummy variable taking the value of 1 if the teacher went to the university or 0 if he or she went to an institute (in Peru, universities usually require 5 years to complete a degree, while pedagogical institutes require around 3 years of education). All teachers in the sample had completed a higher education degree. Finally, given that there has been a large teacher reform in place in Peru for the past few years, we included whether the teacher had applied to be a part of the new program, which is more demanding but also tied to incentives. By having applied voluntarily for the new career, teachers would be showing an interest in leaving the traditional system, which has lower salaries but no evaluations. This variable took the value of 1 if the teacher had applied to enter the new career and 0 if not (we had no data available at the time of data collection on the outcomes of this application).

4.4 Procedures

The School Survey instruments were administered in November of 2011, toward the end of the school year. The PCK test was solved individually by each teacher under fieldworker supervision, with a time limit of 40 min, while the mathematics examination was taken to students in groups with a maximum time of 60 min. Guerrero et al. (2013) report on the school survey procedures, with detailed information about the procedures and preliminary results.

4.5 Methodology

With regard to the analysis, to answer the first research question, we correlated the wealth index of children at age 1 year with our measure of teachers' PCK. For the second question, to estimate the effect of this variable on students' mathematics achievement, we used an ordinary least squares (OLS) regression with cluster variance estimation (Rogers, 1993), which estimates the net association between PCK and achievement. The cluster variance estimation does not have a restriction in the number of observations per cluster (school) and gives accurate standard errors by taking into consideration the correlation among students within each school[9]. This clustering correction was taken into account for both the first and the second question. Finally, to answer the third research question, we ran a regression with teachers as the unit of analysis.

5 Results

Tables 2 and 3 include information about student and teacher characteristics by quartiles of mathematics achievement for the test administered in 2011. Regarding student characteristics, there are more girls in the first (lowest) quartile. The CDA score is lowest in the first quartile and highest in the fourth; in other words, the CDA score by age 5 is predictive of mathematics achievement by age 11. This suggests the importance of intervening early to strengthen

[9] Another option to address the issue of clustering of students is to run a multilevel analysis (Radudenbush & Bryk, 2002), which makes it possible to take into consideration the nested structure of the data and obtain accurate standard errors for each variable included in the model. However, we have an unbalanced number of students in clusters (in many cases one or two students per school) making it inappropriate to use a multilevel approach given that the number of observations to estimate a random component is quite small.

Table 2 Student characteristics by quartile of mathematics achievement

	First quartile ($n = 80$)	Second quartile ($n = 83$)	Third quartile ($n = 71$)	Fourth quartile ($n = 78$)	Total ($n = 312$)
Female (%)	57.5	42.2	52.1	44.9	49.0
Score in the CDA test in 2006	278.3	284.9	303.3	309.8	293.6
Students with indigenous mother tongue (%)	27.5	16.9	11.3	5.1	15.4
Mother has incomplete secondary education or more (%)	27.5	53.0	67.6	75.6	55.5
Wealth index in 2002	0.3	0.4	0.5	0.6	0.5
Wealth index in 2009	0.4	0.5	0.6	0.7	0.5

Source: School Survey—Young Lives

children's mathematics skills. There are more students with indigenous mother tongue in the first quartile and also fewer mothers with incomplete secondary education or less. Finally, the wealth index in 2002 and 2009 is lowest in the first quartile and highest in the fourth. This shows that mathematics achievement can be predicted with socioeconomic variables taken when the child was 1 year old (i.e., wealth index and maternal education).

Table 3 includes information about teacher characteristics. The difference between quartiles with regard to gender or age does not show a clear pattern. Indigenous teachers are more frequent in the first quartile, which corresponds with students' mother tongue as presented above. Teachers with a university degree are found more frequently in the higher quartile. There were no clear patterns with regard to teacher experience and training, and the percentage of application to the new public career is higher in the fourth quartile.

In order to answer our first research question, we used three variables for socioeconomic status: wealth index at age 1 year of the student (i.e., round 1 of Young Lives), wealth index for the same children at age 8 (round 3 of Young Lives), and years of schooling of the mother by the time the child was 1 year of age. The correlations between these and the score of the teacher on our measure of PCK were $r = 0.21$, $r = 0.19$, and $r = 0.28$[10], respectively, all statistically significant ($p < 0.05$). These results reinforce the idea that socioeconomic status is linked with the distribution of teacher quality (as measured in our survey); in the interpretation, it should be considered that as mentioned above, the 5 % richest districts in Peru were eliminated from the sampling design. This would exclude the elite private schools (which charge fees in Peru), so these correlations could be considered an underestimation of what the association could be for the whole country.

With regard to our second research question, we correlated teachers' PCK and student achievement in mathematics and the coefficient was 0.17 ($p < 0.05$). Besides the correlation between continuous scores, in this paper, we explore the importance of having a cutoff score, only above which PCK abilities may be relevant for improving student performance (this is following Marshall et al., 2009, presented above). We decided to use a cutoff above 75 % because it allowed us to differentiate better in this sample (i.e., the number of teachers above the cutoff score was 53 and below 94).

[10] Correlations were calculated by using OLS regression with cluster variance, taking into account the nested structure in the data.

Table 3 Teacher characteristics by quartile of mathematics achievement

	First quartile (*n* = 80)	Second quartile (*n* = 83)	Third quartile (*n* = 71)	Fourth quartile (*n* = 78)	Total (*n* = 312)
Female (%)	52.5	60.2	64.8	65.4	60.6
Age (years)	45.7	44.3	45.0	45.5	45.1
Teachers with indigenous mother tongue (%)	38.8	15.7	12.7	6.4	18.6
Years of teaching experience	19.1	18.4	19.2	19.7	19.1
Teacher went to university (%)	41.3	38.6	53.5	55.1	46.8
In-service training in the last 2 years (scale results)	0.7	0.9	0.7	0.8	0.8
Teacher applied to the new public career (%)	16.3	16.9	23.9	24.4	20.2

Source: School Survey—Young Lives

Thus, for this paper, we gave a score of 1 to teachers who scored at or above the 75 % percentile cutoff score (that is a score of 7 or above) and 0 to those below. In Fig. 1, we present the results of teachers by quartile of student achievement. The gradient is clear up to quartile 3, with students with higher scores being more likely to have a teacher with higher PCK. There is almost no difference between quartiles 3 and 4.

In order to test this association more rigorously, we estimated an OLS regression, controlling for clustering as mentioned above. First, we present the association by using the continuous PCK score. Model 1 in Table 4 shows the unadjusted effect of teachers' PCK on mathematics achievement, which is positive and significant, although the proportion of variance explained is small. Model 2 shows that the coefficient for PCK is non-significant, mostly due to the inclusion of wealth. This variable was highly significant, which is an interesting result in itself: Wealth at home by the time children were 1 year of age predicts not only PCK, as shown above, but also achievement 10 years later. Furthermore, decreases in wealth were also negatively associated with achievement. Rural schools, which tend to be attended by poorer,

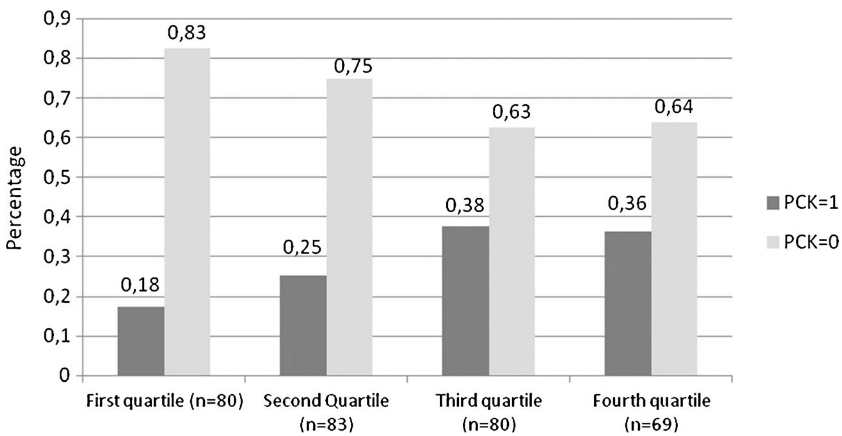

Fig. 1 Percentage of students by level of teacher's PCK and quartiles of mathematics achievement

Table 4 Effect of PCK in mathematic achievement controlling for other variables (standardized coefficients)

Variables	PCK as a continuous score			PCK as a threshold (75 %)		
	M1	M2	M3	M1	M2	M3
Pedagogical content knowledge (PCK)	0.16**	0.00	−0.05	0.19**	0.10+	0.07
Socioeconomic status (SES) in 2002		0.27***	0.25***		0.27***	0.25**
Family wealth decrease over time (2002–2009)		−0.13*	−0.12*		−0.12*	−0.12*
CDA score at 5 years old		0.13*	0.12*		0.12*	0.12*
Mother has incomplete secondary or more		0.13*	0.12*		0.12*	0.11*
Sex (female)		−0.10+	−0.08		−0.09+	−0.07
Age in months		0.06	0.06		0.06	0.06
Indigenous mother tongue		−0.01	−0.01		0.00	0.01
The school is rural		−0.19**	−0.18**		−0.19**	−0.18*
The school is public		−0.05	−0.03		−0.04	−0.04
Teacher variables[a]			Yes			Yes
Number of observations	312	312	312	312	312	312
R-squared	0.03	0.33	0.35	0.03	0.34	0.35

Average variance inflation factor in model 3: 1.42 (for both models)

[a] Teacher variables include sex (female), mother tongue (indigenous), years of experience, teacher went to university, hours of training, and if teacher applied voluntarily to the new career

*** $p < 0.001$, **$p < 0.01$, *$p < 0.05$, $^{+}p < 0.1$

indigenous students, performed significantly lower than urban schools. Preschool abilities and maternal education, the latter a common measure of socioeconomic status, were also significantly associated with achievement. The inclusion of these variables significantly increases the explained variance of the dependent variable. Finally, in model 3, PCK is also non-significant. Model 1 of the second analysis (using PCK as a dummy, with the same threshold of 75 percentile mentioned above) shows the unadjusted effect of teachers' PCK on mathematics achievement, which is positive and significant, as in the previous analysis. The final model (3) includes teacher characteristics, which increase the *R*-squared by only 1 % but made the net effect of PCK on students' achievement no longer significant. This is likely due to the association among some teacher characteristics and PCK presented in Table 4 (however, the levels of multicollinearity were acceptable in the analysis).

Given the results from the last model of Table 4, and related to our third objective, an additional analysis was performed to identify which teacher and school characteristics were associated with higher scores in our measure of teachers' PCK. Given that we used the dichotomous version of the variable, a non-linear model was estimated. Table 5 presents the marginal effects of each variable on PCK. The results show that being a woman or having an indigenous mother tongue reduces the probability of being in the higher scoring category in 21 and 34 %, respectively. Also, every additional year of experience reduces this probability by 3 %. An interesting finding is that neither training nor if a teacher went to university are significant variables. We have no information on what types of contents or skills are emphasized in either of these; this should be an interesting topic for further research.

Table 5 Marginal effects from a logistic regression analysis of teacher and school characteristics on PCK

Characteristics	
Female	−21 %**
Years of experience	−3 %***
Indigenous mother tongue	−34 %***
Went to university	14 %
In-service training	9 %
Teacher applied to the new career	1 %
School is rural	13 %
School is public	6 %
Observations	145
Pseudo R-squared	0.18

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, $^{+}p < 0.1$

6 Discussion

Internationally, there has been an ongoing discussion on the importance of setting goals and measuring both opportunities to learn and achievement for different groups of children[11]. However, often, the data available to look at these simultaneously are scarce, especially for developing countries. In this paper, we have performed such an analysis by using a longitudinal database. More specifically, we have explored the association between early and longitudinal measures of socioeconomic status, teachers' knowledge of content and students (one aspect within the domain of PCK) in fourth grade, and students' achievement in Peru. Below, we briefly present and discuss the results for each of the research questions posed above.

Our first research question referred to the association between socioeconomic status and PCK. The correlation we found was positive and statistically significant between this variable and an index of wealth at the age of 1 and at the age of 8. Our PCK measure also correlated positively and significantly with maternal education. This finding is in line with the description of an unfair educational system, with students from higher socioeconomic levels, from a Spanish-speaking background, ending up with teachers who scored higher in our PCK test. These teachers have higher abilities to understand what students did wrong and what were they thinking in specific mathematics exercises. This should be important in the classrooms to prepare better lessons and be able to help students learn from their mistakes. While the association between socioeconomic status and achievement in schools has been often reported, there are fewer studies showing an association between socioeconomic status and quality of teachers. Furthermore, our study shows that teacher quality is related to student wealth at the ages of 1 and 8 years. The implications of this finding would lead to think of ways of testing abilities of teachers and allocating them in a more equitable way or maybe promoting equality by assigning better qualified teachers to students who are poorer or indigenous, who as shown above showed the poorest results.

The second research question was related to the links between PCK and student achievement. Results showed that students with higher scores are more likely to have a teacher with a higher PCK.

[11] See, for example, recommendations from the Learning Metrics Task Force convened by UNESCO Institute for Statistics and the Center for Universal Education at the Brookings Institution at http://www.brookings.edu/research/reports/2013/09/learning-metrics-task-force-universal-learning.

This association was significant and positive even after controlling for children and family characteristics, but only in the model using PCK as a dichotomous variable. However, in the final model, the PCK effect was no longer significant. By far, the strongest predictor of student achievement by the end of fourth grade was socioeconomic status (the wealth index), captured at age of 1. Furthermore, declines over time in the wealth index, preschool quantitative abilities, as well as maternal education and the location of the school in a rural area were also predictive of achievement, depicting a highly unequal system with regard to mathematics learning by students. This result emphasizes the need for a fair allocation of teachers; there would seem to be a reinforcing circle among students' socioeconomic characteristics, the abilities of their teachers (again, at least as measured through our survey), and the students' scores in a mathematics test. In this paper, we tested PCK as both a continuous variable and a dichotomous variable, to explore whether there is a threshold above which PCK would be more relevant for student achievement. The results suggest that there may be a threshold, as the coefficient for this analysis was more robust than when we used the continuous variable. Again, this seems a relevant topic for further research.

The explanatory power of our PCK measure was nevertheless small. It may be that we need a more complete set of surveys to measure other domains within PCK, for example, subject matter knowledge of teachers, which has been found to predict student achievement in Peru (Metzler & Woessmann, 2012). Additionally, even if teachers show high levels of PCK, they do not necessarily use these skills in classroom interactions, for example, teachers in developing countries are often absent from the classroom (Chaudhury, Hammer, Kremer, Muralidharan, & Rogers, 2006), or they do not give enough feedback to students (Cueto et al., 2006).

While the longitudinal design used in this study does not allow us to control for all potential confounders, we believe that there is enough evidence from this and other studies to develop randomized controlled trials with several aspects of PCK as the focus, carried out in developing countries, and with the goals of increasing scores and reducing inequalities. These designs are usually considered solid for establishing cause and effect results. In targeting teachers, and related to our third objective, we found that in our sample, those who were older, had an indigenous mother tongue, or were female had significantly lower scores than their peers. Again, this has to do at least partially with an unfair educational system, given that as shown above, indigenous teachers tend to be assigned to indigenous, low achieving students.

Currently, in the minds of many researchers and practitioners, teachers' pedagogical knowledge matter for students' learning. However, as noted by Hill et al. (2008), the "conceptualization of the domain is far from straightforward" (p. 396). As mentioned before, several specific constructs of teachers' knowledge have been suggested by researchers, as well as the links among them. This may be one of the reasons behind the difficulties, shown in this and other papers, to measure aspects of PCK with high levels of reliability. The difficulties in defining the construct and its several subcategories (as shown in the literature review section), as well as its perceived importance to explain learning, seem to have spurred the development of a variety of methods, quantitative and qualitative, to measure it. Hill et al. (2008) include knowledge of content and teaching, knowledge of curriculum, and knowledge of content and students under PCK. Our measure in this study aimed at capturing knowledge of content and students or, more operationally, the ability of teachers to identify patterns of mistakes and the reasons why students err in solving mathematics items.

With regard to how to improve the reliability of instruments such as the one presented here, we suggest the inclusion of more items of the same type. Also, given that the test was a voluntary response from teachers, it may be that some of them did not respond with their best effort as the survey had no implications for them. If the test were administered by an official body (e.g., the Ministry of Education) with high stakes (e.g., selection for a teaching post or provision of feedback), teachers would work harder to answer the survey and hence show their full abilities. With regard to validity, further studies would need to be developed showing specifically how teachers implement their knowledge of content and students in their different ways of interacting with the latter. Such studies could also be helpful to refine the survey instruments and develop more comprehensive measures of teachers' abilities.

Research suggests that teachers' PCK may be learned by teachers in formal courses. Teacher preparation and professional development programs could include the learning of the content and students in a systematic and accountable way to ensure that all teachers have this skill and knowledge. For example, Hill and Ball (2004) analyzed the impact of a Mathematics Professional Development Institute in California. The institute lasted between 1 and 3 weeks, depending on the group. The study used a pre-post test design and found a significant increase in teachers' scores of content knowledge for teaching mathematics. The increases were associated with the length of the institute and focus on mathematical analysis, reasoning, and communication. Lim-Teo et al. (2007) found also an increase in mathematics PCK from pre-test to post-test in preservice teachers participating in a Diploma in Education course, after 2 years of training. As in other studies, they presented teachers with pedagogical situations that involved students' responses and asked them to explain their reasoning and plan how to improve their learning.

7 Limitations

As mentioned above, one of the limitations of this study is that the richest 5 % of districts (and thus schools) in Peru were excluded from the sampling design. Given that SES is highly correlated with achievement in Peru and richest children attend private schools with highly qualified teachers, our results are probably an underestimation of the correlation for the whole student population. Second, given the longitudinal nature of the Young Lives data set, it is impossible to establish a cause and effect link because this study does not allow us to control for all potential confounders, although the preschool controls add to the rigor of the analysis. Third, while we used all children in the sample, the number of both teachers and students is small, and this sets a limit to the statistical power; also, the number of students per teacher in the present study is small, which prevents us from using a more rigorous analysis such as a hierarchical model. However, we have used OLS regression with cluster variance that takes into account the nested structure in the data. Fourth, even though both our PCK test and the student test measure number sense and basic operations, the alignment is not one to one and this may have generated an underestimation of the associations.

Finally, although there is no clear consensus about the conceptualization and measurement of PCK, there is no doubt that this kind of teacher knowledge may be learned in pre-service and in-service teacher training. Therefore, this can be an important policy to overcome inequalities in Peruvian educational system and improve student learning.

Acknowledgments Preparation of this study was funded by UK aid from the Department for International Development (DFID) and co-funded by the Netherlands Ministry of Foreign Affairs from 2010 to 2014 and by Irish Aid from 2014 to 2015. The views expressed are those of the authors. They are not necessarily those of, or endorsed by, Young Lives, DFID, or other funders.

Appendix. Sample items of the PCK test

Appendix. Sample Items of the PCK test

6. **Claudia has been correctly solving most math exercises (division); however, she recently started having difficulties. Below are some exercises solved by Claudia**.

A.	B.	C.
413 | 3 3 137 11 9 23 21 2	815 | 2 8 47 015 14 01	626 | 3 6 28 026 24 02

6.1 How do you think Claudia will do in the following exercise if she continues using the same procedure to solve the math exercises (division)?
Mark whit an X only one answer.

928 | 3	a)	☐ 01	She will probably answer it correctly if she uses the same procedure.
	b)	☐ 02	She will probably answer it incorrectly if she uses the same procedure.

Source: Adapted from instrument used in Carnoy and Arends (2012).

References

An, S., Kulm, G., & Wu, Z. (2004). The pedagogical content knowledge of middle school, mathematics teacher in China and the U.S. *Journal of Mathematics Teacher Education, 7*(2), 145–172. doi:10.1023/B:JMTE.0000021943.35739.1c

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. doi:10.3102/0002831209345157

Borman, G. D., & Kimball, S. M. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *The Elementary School Journal, 106*(1), 3–20.

Carnoy, M., & Arends, F. (2012). Explaining mathematics achievement gains in Botswana and South Africa. *Prospects Quarterly Review of Comparative Education, 42*(4), 453–468. doi:10.1007/s11125-012-9246-6

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., & Rogers, H. (2006). Missing in action: Teacher and health worker absence in developing countries. *Journal of Economic Perspectives, 20*(1), 91–116. doi:10.1257/089533006776526058

Conference Board of the Mathematical Sciences. (2001). *The mathematical education of teachers*. Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.
Creemers, B. P. M. (1994). *The effective classroom*. London: Cassell.
Cueto, S. (2013). *Evaluaciones estandarizadas del rendimiento escolar*. Boletín del Consejo Nacional de Educación. Lima, CNE. Retrieved from http://www.cne.gob.pe/images/stories/cne/publicaciones/BOLETIN%20CNE%20setiembre%202013.pdf
Cueto, S., Escobal, J., Penny, M., & Ames, P. (2011). *Tracking disparities: Who gets left behind? Initial Findings from Peru. Round 3 Survey Report* (Young Lives). Oxford: Department of International Development, University of Oxford.
Cueto, S., Leon, J., Guerrero, G., & Munoz, I. (2009). *Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives* (Young Lives Technical Note No. 15). Oxford: University of Oxford.
Cueto, S., Ramirez, C., & Leon, J. (2006). Opportunities to learn and achievement in mathematics in a sample of sixth grade students in Lima, Peru. *Educational Studies in Mathematics, 62*, 25–55. doi:10.1007/s10649-006-7922-2.
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. doi:10.1016/j.tate.2013.03.001
Escobal, J., & Flores, E. (2008). *An assessment of the young lives sampling approach in Peru* (Young Lives Technical Note 3). Retrieved from http://www.younglives.org.uk/files/technical-notes/an-assessment-of-the-young-lives-sampling-approach-in-peru
Grossman, P. L., & Schoenfeld, A. (2005). Teaching subject matter. In L. Darling- Hammond, J. Bransford, P. LePage, K. Hammerness, & H. Duffy (Eds.), *Preparing teachers a changing world: What teachers should learn and be able to do* (pp. 201–231). San Francisco: Jossey-Bass.
Guadalupe, C., León, J., & Cueto, S. (2013). *Charting progress in learning outcomes in Peru using national assessments. Paper commissioned for the EFA Global Monitoring Report 2013/4*. Paris: UNESCO.
Guerrero, G., Leon, J., Rosales, E., Zapata, M., Freire, S., Saldarriaga, V., & Cueto, S. (2013). *Young lives school survey in Peru: Design and initial findings* (Young Lives Working Paper 92). Oxford: University of Oxford.
Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River: Prentice Hall.
Harbison, R., & Hanushek, E. A. (1992). *Educational performance of the poor: Lessons from Rural Northeast Brazil*. New York: Oxford University Press.
Hiebert, J., Morris, A. K., Berk, D., & Jansen, A. (2007). Preparing teachers to learn from teaching. *Journal of Teacher Education, 58*(1), 47–61.
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, 35*(5), 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, 39*(4), 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.
Koh, J. H. L., Chai, C. S., & Tsai, C.-C. (2014). Demographic factors, TPACK constructs, and teachers' perceptions of constructivist-oriented TPACK. *Educational Technology & Society, 17*(1), 185–196.
Koirala, H. P., Davis, M., & Johnson, P. (2008). Development of a performance assessment task and rubric to measure prospective secondary school mathematics teachers' pedagogical content knowledge and skills. *Journal of Mathematics Teacher Education, 11*, 127–138.
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.
Lim-Teo, S. K., Chua, K. G., Cheang, W. K., & Yeo, J. K. (2007). The development of diploma in education student teachers' mathematical pedagogical content knowledge. *International Journal of Mathematics and Science and Mathematics Education, 5*, 237–261. doi:10.1007/s10763-006-9056-5
Marshall, J. H., Chinna, U., Nessay, P., Hok, U. N., Savoeun, V., Tinon, S., & Veasna, M. (2009). Student achievement and education policy in a period of rapid expansion: Assessment data evidence from Cambodia. *International Review of Education, 55*(4), 393–413. doi:10.1007/s11159-009-9133-4
Marshall, J., & 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. doi:10.1007/s11159-012-9276-6
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.
National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston: National Council of Teachers of Mathematics.

Ngo, F. J. (2012). The distribution of pedagogical content knowledge in Cambodia: Gaps and thresholds in math achievement. *Educational Research for Policy and Practice, 12*(2), 81–100.

OECD. (2010). *PISA 2009 Results: Overcoming social background. Equity in learning opportunities and outcomes* (Vol. 2). Retrieved from doi:10.1787/9789264091450-en

OECD. (2013). *PISA 2012 Results: What students know and can do – students performance in mathematics. Reading and science* (Vol. 1). PISA, OECD Publishing. Retrieved from doi:10.1787/9789264201118-en

Park, S., & Oliver, J. S. (2007). Revisiting the conceptualization of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education, 38*(3), 261–184.

Passos, A., Nahara, T., Magaia, F., & Lauchande, C. (2005). The SACMEQ project II in Mozambique: A study of the conditions of schooling and the quality of education. Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) and Ministry of Education and Culture, Mozambique. Retrieved from http://www.sacmeq.org/education-mozambique.htm

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park: Sage Publications.

Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2008). *Can you recognize an effective teacher when you recruit one?* (NBER Working Papers 14485). Cambridge, MA: National Bureau of Economic Research.

Rogers, W. H. (1993). Regression standard errors in clustered samples. *Stata Technical Bulletin, 13*, 19–23.

Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement, 1*, 61–80.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. en Matematica Educativa, 12(2), 251–290.

Sorto, M. A., Marshall, J. H., Luschei, T. F., & Carnoy, M. (2009). Teacher knowledge and teaching in Panama and Costa Rica: A comparative study. *Revista Latinoamericana de Investigación*.

Turnuklu, E. B., & Yesildere, S. (2007). The pedagogical content knowledge in mathematics: Preservice primary mathematics teachers' perspectives in Turkey. *Issues in the Undergraduate Mathematics Preparation of School Teachers: The Journal, 1*, 1–13.

UMC. (2004). *Factores asociados al rendimiento estudiantil. Resultados de la Evaluación Nacional 2001*. Documento de Trabajo n° 9. Lima: Unidad de la Medición de la Calidad Educativa del Ministerio de Educación.

UMC. (2006). *Cómo disminuir la inequidad del sistema educativo peruano y mejorar el rendimiento de sus estudiantes? Factores explicativos de la Evaluación Nacional 2004*. Documento de Trabajo n° 21. Lima: Unidad de la Medición de la Calidad Educativa del Ministerio de Educación.