THE TIMING OF TEACHING PRACTICE: TEACHER KNOWLEDGE AND THE CASE FOR CHILDREN’S MATHEMATICAL THINKING

Erik Jacobson
University of Georgia
ejacobsn@uga.edu

This study explores the hypothesis that experience with children’s mathematical thinking facilitates the development of teachers’ professional knowledge. Multilevel models of a data set representative of US public institutions preparing K-6 teachers from the Teacher Education and Development Survey in Mathematics (Brese & Tatto, 2012) suggested that the timing of some teaching practice prior the last year of the preparation program (i.e., concurrent with or preceding some coursework) had a positive effect on prospective teachers’ content knowledge and was moderated by the length of teaching practice. Teachers’ pedagogical content knowledge was not associated with the timing or length of teaching practice. The findings corroborate prior experimental work at a single institution (Philip et al., 2007) and raise questions about how pedagogical content knowledge develops during teacher education.

Keywords: Teacher Knowledge, Teacher Education-Preservice, Elementary School Education

Recent policy recommendations for teacher preparation in the United States (e.g., National Research Council [NRC], 2010; National Council for the Accreditation of Schools [NCATE], 2010) have focused on the promise of school-based field experiences for producing desired outcomes for preservice teachers. Those recommendations echo an international consensus on the importance of clinical experiences for teacher education (e.g., Musset, 2010; Wang, Coleman, Coley, & Phelps, 2003) and an international trend over the last few decades among teacher educators of increased emphasis on clinical experiences (Maandag, Deinum, Hofman, & Buitink, 2007; Ronfeldt & Reininger, 2012). Policy recommendations call for increased field-based teacher education by scheduling earlier clinical experiences and extending their duration (Goodson, 1993; Villegas-Reimers, 2003). At the same time, descriptive studies of clinical experiences in the United States suggest that they can be poorly aligned with teacher education program goals and that placements in schools can be haphazard, with little university oversight (Wilson, Floden, Ferrini-Mundy, 2001). Some researchers argue that earlier or longer clinical experiences may be ineffective or even detrimental if the quality is poor, for example, by leading to beliefs about mathematics teaching and learning that are inconsistent with university course work (Zeichner & Gore, 1990).

Rather than examining student outcomes directly, this study focuses on teachers’ knowledge outcomes. In recognition that student teaching may not have the same effects across the wide range of content areas and grade levels for which teachers are prepared and in response to urgent calls to find ways to improve mathematics teacher education (e.g., National Mathematics Advisory Panel, 2008), this study focuses on the preparation of elementary (grades K–6) teachers to teach mathematics. I hypothesized that teaching practice that is timed early in the preparation program and thus preceding or concurrent with some content and methods courses will enable otherwise comparable prospective teachers in otherwise comparable programs to develop greater content knowledge and pedagogical content knowledge.

Framework

Teacher Knowledge

Historically, as measures of mathematics teachers’ knowledge have increasingly focused on mathematical knowledge that is used in practice (rather than advanced disciplinary knowledge), the
strength of the observed relationship with student achievement has increased (Hill, Sleep, Lewis, & Ball, 2007). Several measures of teacher knowledge have been framed in light of Shulman’s (1986) notion of pedagogical content knowledge, and these measures are more predictive of student achievement in large-scale studies than covariates such as gender, race, and poverty (e.g., Baumert et al., 2010; Hill et al., 2005).

The Length Of Teaching Practice

There is less empirical support for the recommendations for longer student teaching; the relevant research has been primarily descriptive and frequently lacked adequate controls for selection bias. Some studies have provided evidence for positive effects of extended field experiences on teacher outcomes (e.g., Andrew, 1990, Andrew & Schwab, 1995; Silvernail & Costello, 1983). By contrast, large-scale studies (also lacking adequate controls) have compared teachers completing one versus two semesters of student teaching and found no difference in teaching self-efficacy beliefs (Chambers & Hardy, 2005; Spooner, Flowers, Lambert, & Algozzoine, 2008). Only two studies of which I am aware estimated pseudo-causal effects for student teaching: Boyd, Grossman, Lankford, Loeb, and Wyckoff (2009) used a robust set of controls and found that estimates of the effect of no student teaching on teachers’ value added to student achievement was unstable across models; Ronfeldt and Reinger (2012) used similar controls and concluded that the length of student teaching had no effect on teachers’ preparedness to teach, but that the quality of student teaching had significant positive effects.

Children’s Mathematical Thinking And The Timing of Teaching Practice

Teachers’ experience with children’s mathematical thinking is an important theme in the research on teachers’ knowledge and beliefs about mathematics and learning, and a key marker of quality for field experiences. A randomized experiment (Philipp et al., 2007) compared the change over a semester-long mathematics content course in the beliefs and content knowledge of prospective teachers who were assigned to guided experiences that focused on children’s mathematical thinking with that of prospective teachers assigned to clinical experiences that lacked such a focus. A key design feature was the early timing of the clinical experience to be concurrent with a content course. The authors hypothesized that experience with children’s mathematical thinking would promote the prospective teachers’ development of beliefs and content knowledge, and they found significant differences between the groups with respect to changes in beliefs and greater (although not statistically significant) increases in content knowledge among the students who focused on children’s mathematics. They did not use an instrument to measure pedagogical content knowledge. Silverman and Thompson (2008) have also argued that teachers’ experience and knowledge of children’s thinking is critical for developing mathematical knowledge for teaching, a domain of teacher knowledge that is closely related to pedagogical content knowledge.

Data

This study used data from prospective teacher and teacher preparation program surveys conducted by the Teacher Education and Development Study in Mathematics (TEDS-M, Brese & Tato, 2012). The analytic sample was restricted to US teachers who participated in primary (K–6) programs as identified by TEDS-M. The data set included scales of content knowledge and pedagogical content knowledge, and the U.S. sample was designed to be nationally representative of the public institutions that prepare teachers. The large-scale data sets used in prior research have not included teacher knowledge scales and have been restricted to single school districts (Boyd et al., 2009; Ronfeldt & Reinger, 2012) or states (Goldhaber & Liddle, 2011; Harris & Sass, 2007).

The TEDS-M study distinguished introductory field experiences (e.g., observation) from extended teaching practice. All teacher preparation programs reported both kinds of field experiences, but programs varied with respect to the timing of teaching practice, with many reporting teaching practice prior to the final year of the program. For this study, I defined early teaching
practice as attending a teacher preparation program that involved teaching practice prior to the final year of preparation (i.e., preceding or concurrent with other preparation activities such as content and methods coursework). About 41% of the US institutions in the sample (20 of the 49) scheduled teaching practice before the final year of the teacher preparation program, but only an estimated 35.7% of prospective teachers experienced early teaching practice after adjusting for the survey sample population weights.

The length of teaching practice was operationalized as the total number of contact hours taken as the product of days of extended teaching practice per year and the corresponding annual estimate of the average number of hours per day. To ease interpretation of the analysis, I divided the number of contact hours by 40 to obtain the number of contact weeks. The length of teaching practice ranged from 240 contact hours (6 weeks) to 1224 contact hours (30.6 weeks). There was a clear peak in the distribution around 16 weeks—approximately 1 semester—see Figure 1. Note that some programs achieved this amount of teaching practice over one calendar semester and others spread 16 weeks of teaching practice over one calendar year or more.

![Figure 1: Histogram Of The Length of Total Teaching Practice](image)

The outcomes for the study were operationalized using IRT scales developed for TEDS-M that are well matched to the underlying constructs and well aligned with the research questions and framework of the study. In particular, the TEDS-M scale for pedagogical content knowledge is explicitly based on Shulman’s work (1986).

**Methods**

**Multiple Imputation**

The TEDS-M sample for the United States consisted of 1,119 prospective primary school teachers in 49 concurrent U.S. preparation programs operated by public institutions. Four programs (75 teachers) were missing institutional data on teaching practice and these were excluded from the sample, leaving an analytic sample of 1044 prospective teachers in 45 programs. Of these teachers, 25.3% were missing data on the outcome measures. I used the R program Amelia II (which employs a bootstrap expectation-maximization algorithm; see Honaker, King, & Blackwell, 2011) to create 50 imputed data sets for all 1,119 individuals in the TEDS-M sample using 70 individual- and program-level variables. Observed data from the 75 individuals with missing institutional data and who were excluded from the analytic sample nevertheless contributed to the imputation of missing data for other individuals. The missing data on the variables used for imputation ranged from 0% to 28%. Because of the large number of variables in the data set related to the participants’ prior experience with mathematics, education, and their teacher education programs, I was confident that the distributions of imputed values for the missing teacher knowledge outcome measures were...
Covariates To Mitigate Selection Bias

Selection bias is a concern when comparing teachers in different programs; individual characteristics may be responsible for observed outcomes rather than the teacher education program. Individual and program-level covariates similar to those used in earlier research (Boyd et al., 2009; Ronfeldt & Reininger, 2012) to mitigate bias in estimates of the effects of teaching practice duration and quality are available in the TEDS-M dataset and were used in the proposed analysis. Estimates of the effects of teacher programs can confound selection into the program with program features (Goldhaber & Liddle, 2011; Harris & Sass, 2007), so I also included program-level variables related to selectivity and individual-level high school achievement to control for self-selection bias. The sets of pretreatment covariates at the program and individual levels also included many of the variables used to predict college choice (e.g., Cabrera & La Nasa, 2000) and teacher knowledge (e.g., Hill, 2010). Thus, these covariates likely reduced bias from omitted variables.

Multilevel Modeling With Survey Weights

The complex sampling design of the TEDS-M data was addressed by using multilevel modeling and by incorporating sampling weights into the analysis. For each outcome variable I used the statistical software MPLUS (Version 6.11 for Mac) to estimate a multilevel model (prospective teachers nested within preparation programs) across the 50 imputed data sets with standard errors for testing individual regression coefficients. Weights should not be used without appropriate scaling because unscaled weights can bias estimates (Carle, 2009). Both scaling methods recommended by Carle (cluster sample size and effective cluster sample size) were available in MPLUS, and I used both methods and compared the results. I also ran the analyses without weights. The results across all three methods were very consistent with each other, and I report results from the cluster sample size method.

The multilevel model for this study was adapted from VanderWeele (2008) and is appropriate for estimating neighborhood effects—effects at the program rather than individual level. This model accommodates the expected homogeneity among prospective teachers in the same program (Gelman & Hill, 2007). The first equation in Figure 2 expresses the individual level of the model. The model predicts the outcome $Y$ (prospective teachers’ content knowledge or pedagogical content knowledge) with $i$ indexing individuals and $j$ indexing programs. The matrix $X$ represents the individual-level covariates for individual $i$ in program $j$ (see Table 1); and $e_{ij}$ is the random error term associated with individual $i$ in program $j$. The vector of coefficients $\beta$; are estimated by fitting the model to the observed data; these terms provide estimates of the relationships between these variables and the outcome $Y$. The remaining term in the first equation is the intercept term $\mu$. It represents the average outcome for each program after accounting for differences in the individual level covariates.

$$Y_{ij} = \mu_j + \beta_j X_{ij} + e_{ij}$$

$$\mu_j = \alpha + \gamma_1 T_j + \gamma_2 L_j + \gamma_3 T_j L_j + \beta_j Z_j + u_j$$

$$e_{ij} \sim N(0, \sigma_1); \quad u_j \sim N(0, \sigma_2)$$

Figure 2: Multilevel Model For Neighborhood Effects.

The second equation in Figure 2 predicts $\mu_j$, the mean outcome of program $j$ using program-level variables. The variable $\alpha$ is the overall mean outcome. The binary indicator variable $T_j$ represents whether program $j$ has early teaching practice; the variable $L_j$ represents the length of total teaching practice in program $j$; and the interaction term $T_j L_j$ expresses the possibility of an increase in the mean program outcome for timing and length beyond that accounted for by each variable independently. The matrix $Z$ represents the program level covariates (see Table 1). Finally, $u_{ij}$ is the random error term associated with program $j$. A fitted model provides estimates of the coefficients $\gamma_1$, $\gamma_2$, $\gamma_3$ and of the coefficient vectors $\beta_1$ and $\beta_2$. The last line of Figure 2 indicates the assumption that

---


Articles published in the Proceedings are copyrighted by the authors.
the random error terms at each level are normally distributed.

**Results**

After estimating the full model for both outcomes (content knowledge and pedagogical content knowledge), I compared the results with the corresponding null models (i.e., no individual or program level predictors, just random program level intercepts). The focus of the analysis for this study was on program level differences in outcomes, and I found that the full models explained a large portion of the variance in outcomes at the program level in addition to a moderate portion of the variance in outcomes at the individual level. The full model for content knowledge explained 8% of the individual level variance and 71% of the program level variance that was not explained by the null model, and the full model for pedagogical content knowledge explained 6% of the individual level variance and 74% of the program level variance. Thus, these models explained to a large degree how observed program outcomes differed in relation to program characteristics including the focal variables describing teaching practice.

The estimated coefficients for the full models are reported in Table 1. Overall, the covariates predicting knowledge outcomes appeared to have functioned as expected to mitigate selection bias. Many of the program and individual level covariates in models are significant, and all significant predictors have the expected sign. At the individual level, for example, socioeconomic status and secondary school achievement were statically significant and positive predictors of each outcome as expected. Results also reflected a substantial gender gap in content knowledge, with men \( (n = 111) \) more knowledgeable.

Very different patterns of significant coefficients linked the development of content and pedagogical content knowledge to different features of teacher preparation programs, and provided evidence of different pathways for these two knowledge outcomes. Prospective teachers were asked how many topics related to continuity and functions they had studied in their tertiary coursework (e.g., limits, sequences, derivatives), and the TEDS-M data set included a Rasch scale based on 5 such items. The tertiary mathematics variable was statistically significant and positively related to content knowledge but not significant in the model of pedagogical content knowledge. The TEDS-M data set also included a Rasch rating scale of program coherence based on 6 rating-of-agreement items (e.g., “Later courses in the program built on what was taught in earlier courses in the program.”). The program coherence variable as well as the number of mathematics and mathematics pedagogy classes in the program were statistically significant and positively related to pedagogical content knowledge, but were not significant in the model of content knowledge.

**Table 1. Models of Content Knowledge (CK) and Pedagogical Content Knowledge (PCK)**

<table>
<thead>
<tr>
<th>Term</th>
<th>Model 1: CK B (SE)</th>
<th>Model 2: PCK B (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>510.08 (5.13) ***</td>
<td>537.30 (4.13) ***</td>
</tr>
<tr>
<td>X – Individual level covariates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>-0.04 (0.40)</td>
<td>7.10 (8.02)</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>25.29 (8.38) **</td>
<td>-0.64 (0.36)</td>
</tr>
<tr>
<td>SES (e.g., mothers’ education, number of book)</td>
<td>2.03 (0.99) *</td>
<td>2.28 (1.14) *</td>
</tr>
<tr>
<td>Hindering circumstances (e.g., need loans, need to work)</td>
<td>-0.51 (2.80)</td>
<td>-1.35 (3.06)</td>
</tr>
<tr>
<td>Secondary school achievement</td>
<td>14.40 (2.11) ***</td>
<td>11.96 (2.51) ***</td>
</tr>
<tr>
<td>Tertiary math topics studied - continuity &amp; functions</td>
<td>5.42 (1.83) **</td>
<td>2.87 (2.09)</td>
</tr>
<tr>
<td>Coherence of preparation program</td>
<td>1.19 (1.06)</td>
<td>2.73 (1.04) **</td>
</tr>
</tbody>
</table>


Articles published in the Proceedings are copyrighted by the authors.
I next examined the estimated coefficients of principal interest: those corresponding to the length and timing of teaching practice. The most notable result from Model 1 (content knowledge) pertaining to teaching practice was the significant interaction term for timing and length of teaching practice (see Table 1). The length of the teaching practice moderated the effect of early timing on the teachers’ content knowledge. Using the simple slope method (Preacher, Curran, & Bauer, 2006), I found the region of significance ($\alpha = .05$). Early teaching practice had a statistically significant effect when the length of total teaching practice was less than 13.7 contact weeks or more than 28.1 contact weeks. Very little data were in the upper region of significance, but the lower region of significance was informative because that region included all programs with short teaching practice—more than 30% of the sample. For these programs, the shorter the length of teaching practice in the program, the greater the estimated effect of early timing of teaching practice on content knowledge. The estimated effect ranged between 54 points at 5 weeks of teaching practice ($2 SD$ below the median length) and 14 points at 16 weeks of teaching practice (the median length). Thus, early timing of teaching practice had an average effect size of approximately $.30 SD$ in the region of significance.

By contrast, none of the predictors related to teaching practice in Model 2 (pedagogical content knowledge) were significant, meaning that the preparation programs achieved similar outcomes with respect to pedagogical content knowledge regardless of the features of student teaching after controlling for other covariates. This result was surprising because pedagogical content knowledge—even more than the content knowledge—was hypothesized to develop in the context of teaching practice. A possible explanation is that the preparation programs did little to influence the prospective teachers’ pedagogical content knowledge because the classes and teaching practice had not been designed for that outcome. Given the widely cited research about the deficits in elementary teachers’ mathematical knowledge (e.g., Ma, 1999) and the large number of other content areas that K-6 teachers must be prepared to teach, mathematics teacher educators may spend the available time focused on content knowledge instead.

**Discussion**

This study provided evidence of the critical role that experience with children’s thinking may play in the development of teachers’ knowledge. In particular, the finding that early teaching practice has a significant and positive relationship with teacher knowledge in programs with less than 13.7 contact weeks of teaching practice echoed the relative knowledge gains of prospective teachers who focused on children’s thinking (Philip et al., 2007). One limitation of this study is that the teaching practice was not directly observed, whereas explicit discussion of student thinking was one of the variables that was manipulated experimentally in Philip et al.’s (2007) study. The inferences in the present study are based on the assumption that teaching practice includes opportunities to attend to student thinking. Yet in spite of this limitation, the signal from these data remains clear, and the

---


Articles published in the Proceedings are copyrighted by the authors.
findings confirm (on a national scale) the hypothesized relationship between the experience of children’s thinking and teachers’ content knowledge.

Philip et al. (2007) hypothesized that the early experience with children’s mathematical thinking would have an even stronger effect on pedagogical content knowledge than on content knowledge (pp. 468-469). Instead, the results of this study showed that the timing and length of teaching practice was not related to pedagogical content knowledge, and they challenge mathematics education researchers to provide a clearer theoretical account of how teacher knowledge develops in teacher education programs.

Acknowledgements

This research was supported by a grant from the American Educational Research Association which receives funds for its "AERA Grants Program" from the National Science Foundation under Grant #DRL-0941014. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

References


