Planning and Changing
Vol. 50, No. 1/2, 2021, pp. 37–57

IMPROVING MATH AND SCIENCE INSTRUCTIONAL QUALITY: EVIDENCE FROM A RURAL CONSORTIUM-BASED REFORM INITIATIVE

Rural schools face particular challenges with teacher quality. The Northeast Tennessee College and Career Ready Consortium was a collaborative reform effort among rural schools with a key goal of improving the quality of instruction in math and science to support its efforts to expand academically rigorous courses. This study examines progress made by the Consortium relative to a group of matched comparison schools in improving instructional quality between 2011 and 2014 during the implementation of a federal Investing in Innovation Fund (i3) grant. Based on a difference-in-differences analysis of over 400 classroom observations, we find evidence of broad-based instructional quality gains in the Consortium, particularly among advanced science classes. Although some of the improvement may be attributable to Consortium-specific activities, some improvements may also stem from statewide and nationwide initiatives during this period. We conclude with implications for rural schools in other settings.

Keywords: instructional quality, rural schools, difference-in-differences analysis, classroom observations

Rural schools face particular challenges in ensuring high quality teaching which has important implications given that instructional quality is an important determinant of college readiness. A recent review of rural education studies in Appalachia found that school administrators often have difficulty attracting and retaining high-quality teachers in geographically remote locations, and teachers who are from the region may lack adequate preparation to teach to rigorous standards. Schools often face teacher shortages in STEM-related fields which commonly leads to out-of-field teaching (Kannapel et al., 2015). In addition, maintaining teacher quality is challenging as access to professional development is limited by geographic isolation and lack of necessary staff such as instructional coaches in rural districts (Hansen, 2009; Rude & Brewer, 2003).

Recognizing these challenges, the Northeast Tennessee College and Career Ready Consortium (the Consortium) set a key goal of improving the quality of instruction in math and science to support its efforts to expand academically rigorous courses. Professional development was one of several strategies used by the Consortium to improve college and career readiness in its schools. Other strategies included expanding access to courses through distance and online technology, increasing opportunities for college-level courses through Advanced Placement (AP) and dual enrollment, and providing a college and career counselors team to promote a college-going culture. The Consortium consisted of a network of 15 neighboring Tennessee cities and counties in rural northeast Tennessee comprising 29 high schools working in partnership with five area colleges. The Consortium's activities were supported by an Investing in Innovation Fund (i3) grant to the Niswonger Foundation between 2010/11 and 2014/15.

According to the theory of change underlying Consortium efforts, the instructional quality of courses students take is an important determinant of students' readiness for college and careers. This theory is backed by research suggesting that instructional quality may be the most important school factor influencing student achievement (Darling-Hammond, 2000; Tharp & Gallimore, 1988). One method commonly used to improve instructional quality is to provide teachers with professional development. Although there is not much conclusive evidence about whether professional development affects student learning and achievement (Jacob et al., 2010; Podgursky et al. 2009), professional development provided to Consortium teachers embodied many research-based best practices.

Our study examines the extent to which math and science teachers who received access to additional professional development in the Consortium strengthened their instructional practices relative to teachers in similar non-Consortium schools. We address the following questions:

1) How does the change in average overall instructional quality ratings between baseline and the end of the grant differ between Consortium and comparison schools?

2) In which subscales are there gains between baseline and the end of the grant?

3) How do these findings vary by course subject and level?

This study contributes to a larger body of literature on the effectiveness of teacher professional development in three important ways. First, it focuses on the effects of professional development in rural high schools which tend to face unique challenges in teacher quality. Previous empirical studies in the literature have focused largely on elementary teachers and urban school contexts (Yoon et al., 2007). Second, this study uses a classroom observation instrument to assess the quality of instruction instead of teachers' self-reported data on changes. Many studies of this type use self-reported teacher data, and prior research has shown that observations tend to provide better data on instruction than surveys or teacher logs (e.g., Porter, 2002). Third, many studies of this type tend to be descriptive in nature which limits the internal validity of the results. This study uses a more rigorous identification strategy with a difference-in-differences design to control for treatment and control-group-level fixed unobserverables.

We begin by describing the professional development provided to Consortium teachers under the i3 grant. Next, we review the literature on attributes of effective professional development and describe how these attributes compare to the professional development provided in the Consortium. Then we provide a study overview and context, followed by a description of the data and methods used for the analysis. We conclude with a discussion of how Consortium activities, along with other state and federal initiatives, may have contributed to improvements in instructional quality as well as implications for rural schools in other settings.

Consortium Professional Development for Teachers

Funding from the i3 grant provided three types of professional development opportunities for Consortium teachers: (1) College Board AP workshops and summer institutes, (2) AP summer professional development academies organized by the Niswonger Foundation, and (3) professional development symposia organized by the Niswonger Foundation. The AP workshops and summer institutes were offered by the College Board, the organization that administers the AP program nationwide. These training opportunities focused on specific AP courses and their prerequisites. Topics included aligning classroom instruction to AP course goals, identifying skills assessed on the AP exam and areas where students need more preparation, drafting course syllabi that meet AP curricular requirements, and designing instruction to provide equitable access to students. The trainings also provided structured time for teachers to network with each other and exchange ideas about teaching AP courses. Grant funding also made the opportunity available to current AP teachers as well as those potentially interested in teaching these courses in the future or applying AP instructional strategies to their other classes.

The second form of professional development provided through the i3 grant was a series of summer "AP Academies" for current and future AP teachers. Organized by the Niswonger Foundation with input from Consortium school teachers, the summer academies ranged in length from one to five days. All academies were led by local AP teachers from Consortium schools and were designed to improve the quality of instruction in existing AP courses. Academy topics included: ideas for ongoing test preparation, new technologies available for laboratory investigations, strategies for preparing and assessing Socratic seminars, and strategies for helping students answer free-response questions effectively on AP exams.

The third form of professional development provided through the i3 grant consisted of annual symposia that were made available to all teachers in Consortium schools. Teachers participated in large group training sessions followed by smaller break-out groups among teachers in the same subject areas. For example, one symposium centered around providing personalized learning by tailoring learning experiences to the individual needs of students. The state's Education Commissioner gave a keynote address that introduced a new "digital dashboard" being implemented at schools statewide to provide real-time data to identify struggling students.

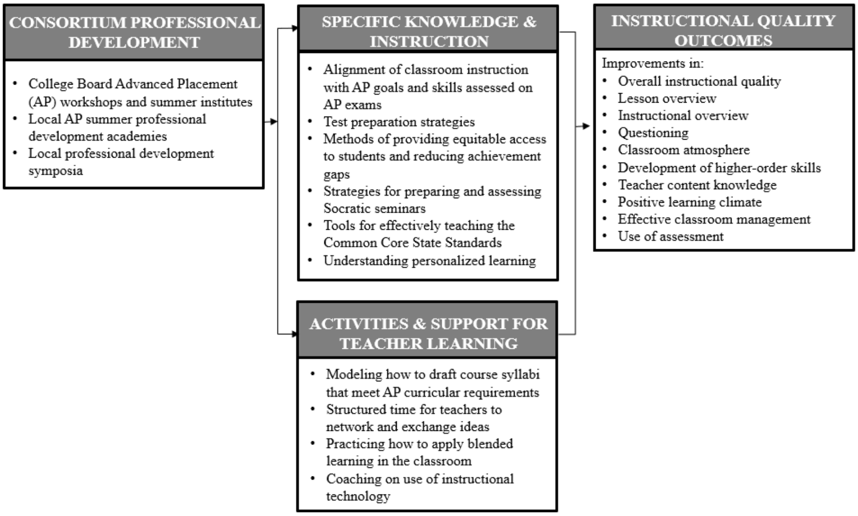
The symposia were also coupled with "hands-on" collaborative days that provided additional opportunities for teachers to practice the techniques learned from the symposia. For the symposium focused on personalized learning, teachers participating in the hands-on day learned how to apply blended learning in the classroom by representatives in the Iredale School District in North Carolina which is considered a national leader in the use of technology. Participants rotated among different practice stations and also developed plans and activities for blended lessons that they could implement in their own classrooms.

Theoretical Framework

According to Wayne and colleagues (2008), the impacts of professional development are moderated through a theory of instruction and a theory of teacher change. Under the theory of instruction, professional development conveys specific knowledge and instruction to teachers about how to improve student achievement. This commonly refers to a focus on instructional practices such as a phonics-based approach to reading. In contrast, the theory of teacher change refers to elements of activities that teachers participate in during professional development which are intended to support teacher learning. For example, a coaching component in a professional development program may provide teachers with individualized guidance and feedback on how to improve instruction. This conflux of influences between the type of content conveyed and the practices engaged in by teachers during professional development makes it difficult to distinguish the specific mechanisms through which professional development may improve teaching in the classroom. Yet taken together, they provide a better understanding of the impact of the *package* of a given professional development intervention (Wayne et al., 2008). Figure 1 illustrates our logic model showing how the specific knowledge and instruction as well as the activities and support for teacher learning from the Consortium professional development are intended to improve instructional quality.

Figure 1

Logic Model for how the Consortium Professional Development is Intended to Improve Instructional Quality Outcomes



Literature Review

Although the literature on the effects of professional development on student performance is inconclusive (e.g. Jacob et al. 2010), there is a growing body of literature, including several longitudinal studies, that examines which approaches to and characteristics of professional development are associated with positive changes in teachers' instructional practice. Garet et al. (2001) identified a framework of features of effective professional development practices that have been supported by subsequent research (e.g., Desimone et al., 2002; Yoon et al., 2007). The framework includes both core features of professional development content and structural features. Core features include a focus on content and related pedagogical knowledge and skills, an active learning approach, and coherence with other instructional initiatives. Among the structural features, longer-term, sustained professional development is associated with changes in practice. Additional research funded by the U.S. Department of Education defines high-quality professional development as sustained and content focused, aligned with state learning standards, and focused on developing an understanding of "scientifically proven" instructional techniques (Yoon et al., 2007).

The Consortium is a collaborative effort intended to improve both the availability and quality of training opportunities for rural teachers. The Consortium professional development activities are conceptually well-aligned with many of the core features associated with effective professional development. The content covered in the grant-funded professional development was primarily subject-specific and included both pedagogical techniques and alignment with state standards. Many of the AP academies and hands-on days focused on active learning to provide teachers with opportunities to put into practice the skills they had learned. The structure of the professional development was designed to encourage collective participation, both by providing symposia to all teachers Consortium-wide and by making AP training opportunities available to both AP and non-AP teachers.

Study Overview and Context

In this study, we use evidence from classroom observations to examine whether math and science instructional quality in Consortium schools improved between the beginning of grant activities in 2010/11 and the end of the grant in 2014/15. The two sets of observations were conducted at Consortium schools and a matched group of comparison schools. The comparison schools are a group of 29 non-Consortium Tennessee high schools selected at the beginning of the grant using propensity score matching based on a number of criteria, including student demographics, baseline academic performance, school resources, community characteristics, and availability of AP and career and technical education courses. After matching, there were no statistically significant differences between the Consortium and comparison schools on any of these characteristics (Mokher, Lee, & Sun, 2019).

Teachers in the comparison schools participated in business-as-usual conditions for professional development. Tennessee teachers must complete 60 professional development points (equivalent to one clock hour of professional learning) during the 10-year period for professional licensure (Tennessee Department of Education, 2019). Comparison group teachers did not participate in any of the Consortium's local AP summer academies or professional development symposium. Some comparison teachers may have had the opportunity to attend a national AP workshop through the College Board, although we anticipate that this occurred relatively infrequently since the grant did not provide any funding to pay for these workshops or associated travel expenses in the comparison schools.

Data and Methods

To measure and better understand potential changes in instructional quality under the Consortium, we used the Leadership by Design (LBD) classroom observation instrument developed by Briarwood Associates. This instrument has been widely used in Tennessee and elsewhere; classroom observation data have been collected using the LBD instrument for more than 3,000 teachers in over 250 elementary, middle, and high schools in seven different states (e.g., Mokher et al., 2018; Tassell et al., 2012). Projects using the LBD include work funded by the U.S. Department of Education and the National Science Foundation. The LBD also has been adopted by the National Science Teachers Association as a program improvement tool to help assess and improve the quality of instruction in middle and high school classrooms.

The LBD is a comprehensive instrument with which trained observers who are subject matter experts measure the quality of a classroom's instructional practices and capture information about the classroom setting. Using the LBD classroom observation instrument, observers collect descriptive data during classroom observations lasting 45 to 90 minutes. The rubric itself consists of 33 elements spanning nine dimensions: lesson overview, instructional overview, questioning, classroom atmosphere, concept development, teacher's content knowledge, learning climate, classroom management, and assessments (see Table 1). After an observation, an LBD Classroom Observation Rubric is used to assign numeric scores to the observational data. The rubric consists of nine instruction-related subscales, plus an overall rating. Each element is rated on a scale of 1 (low) to 5 (high).

The observer also provides an overall rating of instructional quality on the same five-point scale. This overall rating is an independent rating of instructional quality not simply an average of the nine subscale ratings. The overall rating takes into account the observer's general assessment of classroom instructional quality including the effectiveness of instruction, the degree of alignment with objectives and standards for the course being observed, the level of student engagement, and the value of instruction in developing students' higher-order thinking skills. Before each observation, the teacher was asked to provide a lesson plan that described his or her objectives for the lesson and the standards that would be covered during the lesson. The observer then assessed whether the lesson met the objectives and was aligned with the standards. Observers were required to write comments justifying their overall rating.

The Classroom Observation Process

All observers were experienced math or science teachers who had used the LBD instrument in previous studies. The observers conducted two sets of classroom observations in math and science in each of the 29 Consortium schools and in each of the 28 comparison schools at the beginning of the grant. The schools were informed of the visits beforehand and chose the classrooms to be observed. A mix of regular and advanced courses (including AP, International Baccalaureate (IB), honors, and other higher-level courses) were chosen for observations. A limitation of the study is that each classroom was observed only once per visit. However, observers visited two classrooms per subject area in each school so the school's rating was not based on a single teacher observation.

A second set of classroom observations was conducted at the Consortium and comparison schools in the spring or fall of 2014 near the end of the grant. Whenever possible, the same teachers at baseline were observed again, but if the same teacher was no longer teaching at the school, the principal selected another teacher from the same course subject and level. The same teacher was observed in 73% of the cases.

A total of 442 observations were conducted over the two observation periods—224 at baseline and 218 at the end of the grant period. The observations were split about equally between math (N=222) and science classrooms (N=220), and between Consortium schools (N=227) and comparison schools (N=215). About one-quarter of the observations (N=115) were of advanced classes, and the rest (N=327) were of regular classes. There are a small number of missing observations due to unforeseen circumstances. For example, one school was closed on the planned observation date due to damage from a tornado, and it was not possible to reschedule before the end of the school year.

Table 1

Subscales and Overall Rating Definitions for the Leadership by Design Instruments

Subscale	Definition
Lesson overview	Combines ratings of the quality of lesson objectives, use of instructional resources, content delivery, placement in instructional sequence, and seating arrangement for the lesson.
Instructional overview	Includes measures of student focus, instructional strategies, and awareness of student needs.
Questioning	Combines quality of questions, depth and breadth of participation in discussion, use of target-centered questions, and feedback to responses.
Classroom atmosphere	Integrates ratings of student involvement, classroom management, and classroom culture.
Development of higher-order skills	Combines amount and level of student investigation that takes place with an assessment of the extent to which students' scientific sklls are being developed.
Teacher content knowledge	Combines ratings on quality of communcation, connecting content to life experinces, use of strategies appropriate to content, and ability to present lesson content from various perspectives.
Positive learning climate	Integrates ratings on communicating high expectations, establishing a positive learning environment, valuing and supporting diversity, fostering mutural respect between teacher and students, and providing a safe environment.
Effective classroom management	Includes measures of the extent to which instruction is based on an accurate assessment of student needs; effective use of time, space, and materials; and instruction that facilitates higher-order thinking.
Use of assessment	Combine ratings of alignment of assessment with learning objectives, use of variety of formative and summative assessments, and degree to which the classroom accommodates diverse learning needs.
Overall rating	Instruction was of high quality and effective for all students; evidence that instruction was based on clearly defined objectives that were fully aligned with standards; all students were engaged in activities requiring higher-level thinking skills.

Research Questions and Analysis Plan

Insight into the effects that i3 grant activities have on the instructional quality of Consortium classrooms is a key goal of this analysis. Gaining this insight is complicated because other concurrent influences, such as state Race to the Top initiatives, may affect instructional quality at all schools in the state. Additionally, all teachers may become more effective over time as they make strides in their teaching practices. An important aspect of the analysis, therefore, is to compare ratings between Consortium schools and comparison schools identified through propensity score matching. This comparison accounts for such statewide and experiential changes—that is, the comparison school ratings give us an idea of the pattern we might have observed in the Consortium school ratings in the absence of i3 grant activities.

Analyses are conducted using a two-period panel of observations collected before and after the implementation of Consortium activities. Our estimate of the effect of the intervention on instructional quality is a difference-in-differences estimator that is expressed as:

$$\widehat{\delta}_1 = (quality_{post,i} - quality_{post,c}) - (quality_{pre,i} - quality_{pre,c}),$$

where δ_1 is the difference pre- and post- grant in the average difference of instructional quality between the intervention (i) and comparison (c) schools. This approach strengthens the causal inferences of our analyses because it controls for unobserved, fixed, group-specific characteristics. A limitation of this approach is that we do not have individual teacher-level data on participation in professional development. This means the analyses provide estimates of the intent-to-treat impacts through the provision of professional developmental opportunities in the Consortium rather than the treatment-on-the-treated for individual participants.

To test whether δ_1 is statistically significant from zero, we ran a regression analysis where the model is estimated for observation *o* in school *s* as:

$$quality_{os} = \beta_0 + \delta_o post_o + o\ \beta_1 t_{os} + \delta_1 post * t_{os} + \mu_{os}.$$

The intercept, β_0, is the average instructional quality in the intervention and comparison schools in the period prior to the grant. The parameter δ_o captures changes in all instructional quality ratings in the intervention and comparison schools pre- and post- grant. The coefficient β_1 measures the effect of other interventions not due to the grant. This estimation strategy helps to disentangle the effects of the intervention from other changes that occur over time. The parameter of interest is on the interaction term, where δ_1 measures the difference in instructional quality due to the intervention, provided we assume that both intervention and comparison schools did not experience changes in instructional quality at different rates for other reasons. The error term is μ_{os} which is clustered by school. Models are estimated overall as well as separately for each subject area.

Findings

Overall Rating of Classroom Instructional Quality

We find that both Consortium and comparison schools experienced gains in overall instructional quality over time, but the gains were greater among the Consortium schools. The coefficient for the post-treatment variable indicates that all schools experienced an increase in instructional ratings of 0.260 on a five-point scale before and after the i3 grant (see Table 2, pages 52-53). However, ratings in the Consortium schools increased by an additional 0.217 point. This suggests that instructional quality increased about twice as much over time in the Consortium schools relative to the comparison schools.

The models disaggregated by course level and subject area provide further insight into where the greatest gains occurred in overall instructional quality. We find that the Consortium schools demonstrated the greatest growth in advanced courses while there were no statistically significant differences between Consortium and comparison schools on ratings gains in regular courses. Overall instructional quality ratings in advanced courses increased 0.463 point more in the Consortium schools relative to the comparison schools. These gains were most prevalent in advanced science courses where instructional ratings increased by 0.577 point more in Consortium schools before and after the grant.

Instructional Quality Ratings by Subscale

We find differences in the results for the subscale instructional ratings by course type and subject area. In math, both the Consortium and comparison schools experienced gains over time in all of the subscales except classroom atmosphere (see Table 3, pages 54-55). These gains ranged from 0.262 point (positive learning climate) to 0.725 (use of assessment) on a five-point scale. However, the gains in Consortium schools were no greater than in the comparison schools. In fact, gains were slightly lower in the Consortium schools in regular math classes for the instructional overview subscale. Results were similar among Consortium and comparison schools in advanced math courses.

Among all science classes, there were similar gains in instructional quality in both intervention and comparison schools for five subscales (instructional overview, questioning, teacher content knowledge, positive learning climate, and effective classroom management) (see Table 4, pages 56-57). The one area where the Consortium schools outperformed the comparison schools was the subscale for development of higher-order skills.

The results disaggregated by course type indicate that changes over time in instructional subscales were similar between Consortium and comparison schools in regular science classes. Yet, the Consortium schools outperformed comparison schools on four of the nine subscales in advanced science courses (classroom atmosphere, positive learning climate, effective classroom management, and use of assessment). Almost all of these gains were greater than a half-point and were as large as 1.165 points, these findings indicate that the greatest gains in instructional quality over time occurred among the advanced science courses in the Consortium.

Discussion

Evidence from our evaluation of classroom observation data indicates that instructional quality in math and science improved over time in both Consortium and comparison schools. Gains tended to be greater in Consortium schools, particularly among advanced science classes. One important change that may have influenced instructional quality in both Consortium and comparison schools was the statewide introduction of CCSS in math in 2010. Common Core was intended to change math instruction by promoting a greater focus on fewer topics and instituting a more coherent progression of topics from grade to grade (Gewertz, 2015).

Another set of changes that may have influenced both groups of schools simultaneously stem from the national Race to the Top program, in which Tennessee participated beginning in spring 2010. Race to the Top required states to make changes in four core areas: establishing high standards, developing and supporting effective teachers and leaders, creating data systems and using technology to enhance instruction, and turning around low-performing schools. Examples of activities that may have influenced instructional quality include training for principals on how to observe classroom practices and provide feedback, training sessions on new statewide academic standards led by high-performing teachers, and development of school action plans for the standards transition by teams of educators (U.S. Department of Education, 2015).

In addition to these larger state and national changes, Consortium-specific changes may have also contributed to changes in instructional quality over time. Several different types of professional development were provided with i3 grant funding; it is not possible to isolate the effects of each. However, because much of the professional development focused on AP, it is not surprising that the largest effects were found among the advanced courses. Even though teachers of non-AP courses were invited to participate in the Consortium's AP training, they participated at lower rates than teachers of AP courses.

Taken together, our findings suggest that improvements in instructional quality were likely achieved in northeast Tennessee through a combination of national, state, and local initiatives. Most of these involved providing educators with greater access to professional development to improve the rigor of coursework, either through new state standards or through the AP program. Given the number of different changes occurring simultaneously, it would be difficult to replicate this finding in another setting. However, this study still provides several implications for improving instructional quality in rural settings.

First, rural teachers who are from the region may lack adequate preparation to teach to rigorous standards and may also be teaching out-of-their-field, particularly in STEM-fields where there tend to be greater shortages in rural areas (Kannapel et al., 2015). The broader literature on characteristics of effective professional development indicates that content should be subject-specific and aligned with state standards (e.g., Smylie et al., 2001; Yoon et al., 2007). Interviews with teachers and administrators indicated that some of the most influential professional development they received focused on understanding and implementing the CCSS (Pearson, Carr, & Miller, 2015). More of this type of training could be particularly beneficial to teachers in other rural settings.

Second, students in rural schools often have less access to rigorous courses than do students in urban or suburban schools (Anderson & Chang 2011; Levin, 2007). Expanding access to rigorous courses may be particularly important for rural schools such as those in the Consortium which historically may have had limited availability of such courses. In order to increase these types of course offerings, schools should consider increasing the number of AP-certified teachers, providing additional professional development related to AP to improve the instructional quality of these courses, and also allowing non-AP teachers to participate in these types of training opportunities so that similar rigorous strategies can be applied schoolwide to better prepare students for advanced course taking.

Finally, it may be difficult to maintain teacher quality as access to professional development is limited by geographic isolation and lack of necessary staff in rural districts (Hansen, 2009; Rude & Brewer, 2003). One way to address this problem is by using a Consortium-based approach that pools resources from a group of surrounding districts. For example, in the Consortium funded through this i3 grant, administrators identified teachers with the highest AP English exam pass rates to share strategies for helping students to pass the AP exam with all other AP English teachers in the region. This type of collaborative approach may allow districts to provide more opportunities than they may be able to otherwise on their own increasing the pool of qualified instructors to draw from.

Table 2

The Effect of the Consortium on Overall Classroom Observation Ratings, by Subject Area

	All						Regular	
Subject	**Post**		**Treat**		**PostXTreat**		**Post**	
Math and Science	0.260 (0.105)	*	-0.832 (0.128)	**	0.217 (0.142)	~	0.225 (0.115)	~
Math only	0.370 (0.118)	**	-0.174 0.128		0.264 (0.181)		0.351 (0.123)	**
Science only	0.143 (0.134)		-0.174 (0.128)		0.176 (0.178)		0.065 (0.180)	

Table 2 (cont.)

Regular (cont.)		Advanced					
Treat	PostXTreat	Post		Treat	PostXTreat		
-0.174	0.115	0.365	*	-0.174	0.463	~
(0.128)	(0.169)	(0.183)		(0.128)	(0.240)	
-0.174	0.208	0.453		-0.174	0.308	
0.128	(0.194)	(0.352)		(0.128)	(0.439)	
-0.174	0.045	0.309	~	-0.174	0.577	*
(0.128)	(0.213)	(0.180)		(0.128)	(0.277)	

Note. These results are based on the difference-in-differences specification described in the text. “Post” refers to the classroom observation period at the end of the i3 grant, “Treat” refers to teachers in schools participating in the Consortium professional development activities, and “PostXTreat” is an interaction term for the difference in instructional quality in the post-grant period due to the intervention. N=422 for math and science, N=222 for math only, and N=220 for science only. Standard errors (in parentheses) are clustered at the school level. Coefficients are statistically significant at the ~p<.10, *p<.05, and **p<0.1 levels.

Table 3

The Effect of the Consortium on Math Classroom Observation Subscale Ratings

	All math						Regular math	
Subscale	**Post**		**Treat**		**PostXTreat**		**Post**	
Lesson overview	0.513	**	-0.039		-0.224		0.498	**
	(0.100)		(0.109)		(0.161)		(0.096)	
Instructional overview	0.699	**	0.049		-0.354	~	0.719	**
	(0.119)		(0.144)		(0.191)		(0.119)	
Questioning	0.658	**	-0.089		-0.324		0.644	**
	(0.136)		(0.150)		(0.228)		(0.143)	
Classroom atmosphere	0.127		-0.071		0.075		0.125	
	(0.095)		(0.124)		(0.167)		(0.110)	
Development of higher-order skills	0.679	**	-0.095		-0.128		0.650	**
	(0.130)		(0.147)		(0.228)		(0.133)	
Teacher content knowledge	0.387	**	0.091		-0.213		0.375	**
	(0.128)		(0.124)		(0.206)		(0.133)	
Positive learning climate	0.262	**	-0.041		0.056		0.236	*
	(0.099)		(0.113)		(0.146)		(0.096)	
Effective classroom management	0.428	**	-0.033		0.059		0.428	**
	(0.118)		(0.144)		(0.173)		(0.112)	
Use of assessment	0.725	**	-0.234		-0.060		0.753	**
	(0.195)		(0.201)		(0.270)		(0.159)	

Table 3 (cont.)

Regular math (cont.)			Advanced math				
Treat	**PostXTreat**		**Post**		**Treat**	**PostXTreat**	
-0.039	-0.213		0.576	*	-0.039	-0.321	
(0.109)	(0.155)		(0.283)		(0.109)	(0.386)	
0.049	-0.437	*	0.614	~	-0.049	-0.104	
(0.144)	(0.204)		(0.357)		(0.144)	(0.446)	
-0.089	-.406		0.171	~	-0.089	-0.153	
(0.150)	(0.246)		(0.373)		(0.151)	(0.464)	
-0.071	0.014		0.137		-0.071	0.207	
(0.124)	(0.181)		(0.284)		(0.124)	(0.373)	
-0.095	-0.239		0.805	~	0.095	0.045	
(0.124)	(0.218)		(0.322)		(0.124)	(0.430)	
0.091	-0.227		0.439		0.091	-0.257	
(0.124)	(0.218)		(0.322)		(0.124)	(0.430)	
-0.041	0.038		0.375		-0.041	0.021	
(0.113)	(0.149)		(0.230)		(0.114)	(0.291)	
-0.033	0.019		0.426		-0.033	0.166	
(0.144)	(0.198)		(0.417)		(0.144)	(0.452)	
-0.234	-0.104		0.606		-0.234	0.056	
(0.201)	(0.277)		(0.496)		(0.201)	(0.569)	

Note. These results are based on the difference-in-differences specification described in the text. “Post” refers to the classroom observation period at the end of the i3 grant, “Treat” refers to teachers in schools participating in the Consortium professional development activities, and “PostXTreat” is an interaction term for the difference in instructional quality in the post-grant period due to the intervention. N=222 for all math, N=164 for regular math, and N=57 for advanced math. Standard errors (in parentheses) are clustered at the school level. Coefficients are statistically significant at the ~p<.10, *p<.05, and **p<0.1 levels.

Table 4

The Effect of the Consortium on Science Classroom Observation Subscale Ratings

	All Science						Regular science	
Subscale	**Post**		**Treat**	**PostXTreat**			**Post**	
Lesson Overview	0.173	**	-0.039	0.023			0.077	
	(0.124)		(0.109)	(0.188)			(0.189)	
Instructional Overview	0.500	**	0.049	-0.150			0.402	*
	(0.140)		(0.144)	(0.200)			(0.189)	
Questioning	0.504	**	-0.089	-0.100			0.462	*
	(0.153)		(0.150)	(0.231)			(0.204)	
Classroom Atmosphere	0.075		-0.071	0.113			-0.030	
	(0.134)		(0.124)	(0.193)			(0.180)	
Development of Higher-order Skills	-0.235	~	-0.095	0.348	~		-0.412	~
	(0.140)		(0.147)	(0.213)			(0.210)	
Teacher Content Knowledge	0.471	**	0.091	-0.033			0.396	*
	(0.141)		(0.124)	(0.207)			(0.185)	
Positive Learning Climate	0.234	*	-0.041	0.058			0.196	
	(0.104)		(0.113)	(0.171)			(0.140)	
Effective Classroom Management	0.384	**	-0.033	0.058			0.357	~
	(0.145)		(0.144)	(0.213)			(0.194)	
Use of Assessment	0.129		-0.234	0.362			0.306	
	(0.234)		(0.201)	(0.300)			(0.258)	

Table 4 (cont.)

Regular Science (cont.)		Advanced Science				
Treat	PostXTreat	Post		Treat	PostXTreat	
-0.039	-0.026	0.365	~	-0.039	0.209	
(0.109)	(0.225)	(0.188)		(0.109)	(0.259)	
0.049	-0.284	0.690	**	0.049	0.284	
(0.144)	(0.239)	(0.176)		(0.144)	(0.268)	
-0.089	-0.278	0.588	*	-0.089	0.398	
(0.150)	(0.276)	(0.222)		(0.151)	(0.332)	
-0.071	0.016	0.284		-0.071	0.493	*
(0.124)	(0.251)	(0.223)		(0.124)	(0.284)	
-0.041	-0.006	0.311	*	-0.041	0.306	~
(0.113)	(0.207)	(0.118)		(0.114)	(0.168)	
0.091	-0.138	0.621	**	0.091	0.233	
(0.124)	(0.251)	(0.223)		(0.124)	(0.284)	
-0.041	-0.006	0.311	*	-0.041	0.306	~
(0.113)	(0.207)	(0.118)		(0.114)	(0.168)	
-0.33	-0.086	0.440	*	-0.033	0.501	
(0.144)	(0.257)	(0.202)		0.144	(0.301)	~
-0.234	0.000	-0.224		-0.234	1.165	*
(0.201)	(0.317)	(0.251)		(0.201)	(0.404)	

Note. These results are based on the difference-in-differences specification described in the text. “Post” refers to the classroom observation period at the end of the i3 grant, “Treat” refers to teachers in schools participating in the Consortium professional development activities, and “PostXTreat” is an interaction term for the difference in instructional quality in the post-grant period due to the intervention. N=220 for all science, N=161 for regular science, and N=58 for advanced science. Standard errors (in parentheses) are clustered at the school level. Coefficients are statistically significant at the ~p<.10, *p<.05, and **p<0.1 levels.

References

Anderson, R., & Chang, B. (2011). Mathematics course-taking in rural high schools. *Journal of Research in Rural Education*, *26*(1): 1–10.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, *8*(1).

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81-112.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945.

Gewertz, C. (2015). The Common Core explained. *Education Week*, September 28, 2015.

Hansen, J. W. (2009). *Professional development in rural public schools: A review of literature* [Unpublished doctoral dissertation]. Northern Michigan University.

Jacob, B. A., Lefgren, L., & Sims, D. P. (2010). The persistence of teacher-induced learning. *Journal of Human Resources*, *45*(4), 915-943.

Kannapel, P., Flory, M., Cramer, E., & Carr, R. (2015). *Appalachia rising*. http://bit.ly/1LP4fHV

Levin, H. (2007). What are the mechanisms of high-poverty disadvantages? On the relationship between poverty and curriculum. *North Carolina Law Review*, *85*(1381), 1–38.

Mokher, C.G., Cavalluzzo, L., & Henderson, S. (2018). Examining teacher's instructional practices as they progress through the National Board Certification process. In D. Polly, A. Good, T. Petty, & M. Putman (Eds.). H*andbook of research on professional development for quality teaching and learning* (pp. 464-488). IGI Global.

Mokher, C.G., Lee, S., & Sun, C. (2019). Evaluating innovations for improving college and career readiness in rural schools: Final validation results from an i3 grant. *Research in the Schools*, *26*(1):48-63.

Pearson, J., Carr, R., & Miller, A. (2015). *Promising practices, challenges, and sustainability efforts in the Northeast Tennessee i3 Consortium: Site visit report*. CNA.

Podgursky, M., Springer, M.G. & Hutton, R. (2009). Teacher training and preparation in the United States. In P. Peterson, E. Baker, and B. McGaw (Eds.), *International encyclopedia of education*. Elsevier.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, *31*(7), 3-14.

Rude, H. A., & Brewer, R. D. (2003). Assessment of professional development systems: Improving rural special education services. *Rural Spe-*

cial Education Quarterly, *22*(4), 20–28.

Smylie, M. A., Allensworth, E., Greenberg, R. C., Harris, R., & Luppescu, S. (2001). *Teacher professional development in Chicago: Supporting effective practice*. Consortium on Chicago School Research. https://consortium.uchicago.edu/publications/teacher-professional-development-chicago-supporting-effective-practice

Tassell, J. L., McDaniel, K., Johnson, H., Norman, A., Pankratz, R., & Tyler, R. (2012). EXPLORE performance in mathematics and science. *International Journal of Innovation in Science and Mathematics Education*, *20*(1).

Tennessee Department of Education (2019). *FAQ for educators: Professional development points (PDPs)*. https://www.tn.gov/education/licensing/licensing-faq/pdp-faq-for-educators.html

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. Cambridge University Press.

U.S. Department of Education, Office of Elementary and Secondary Education (2015). *Innovation in America's schools under Race to the Top*. http://www2.ed.gov/programs/racetothetop/rttfinalrptfull.pdf

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, *37*(8), 469-479.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–No. 033). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf

Christine Mokher is an associate professor at Florida State University.

Juliana Pearson is an associate research analyst at the Center for Naval Analyses (CNA) in Virginia.

Thomas Geraghty is a research scienctist at the Center for Naval Analyses (CNA) in Virginia.