Teacher Content Knowledge in Indian Secondary Schools and Its Relationship with Student Learning

By SABRIN BEG, ANNE FITZPATRICK, JASON T. KERWIN, ADRIENNE LUCAS, AND KHANDKER WAHEDUR RAHMAN*

Teachers need content knowledge to teach that content to students. But how much do teachers actually know? Research from Africa suggests that many know very little, and that this is bad for their students (Bold et al. 2017). Whether this pattern holds in other regions is an open question. We provide the first evidence on teacher content knowledge in India, using data we collected in secondary schools in the state of Odisha. We document three key facts. First, teachers have very limited knowledge of the subjects they are teaching, and this knowledge varies widely across teachers and subjects. Second, teacher basic demographic measures are not strong predictors of teacher knowledge. Third, our results suggest that a remedial instruction intervention changed the relationship between teacher knowledge and student learning: in schools with the remedial instruction, higher teacher knowledge leads to higher student scores, while in control schools teacher knowledge and student test scores are unrelated.

Secondary schools in Odisha, India are an ideal setting for measuring teacher content knowledge and its effects on student learning. The state reflects the worldwide "learning crisis", where learning levels trail significantly behind national curriculum levels despite high enrollment (World Bank, 2018). Secondary schools also provide an opportunity to test content knowledge of teachers in secondary school schools where subject knowledge is more demanding.

I. Empirical Strategy

We first test for the correlation between teacher characteristics and teacher content knowledge. Then, we test the relationship between that content knowledge and student test score growth. Formally, we estimate the following regression:

$$Tea\_Knowledge_j = \alpha + \beta X_j + \varepsilon_j \tag{1}$$

where $Tea\_Knowledge_j$ indicates the percent correct on the content knowledge exam for teacher j and X is a vector of teacher-specific covariates that are often thought to be associated with teacher content knowledge. We estimate this regression separately for knowledge of English content by English teachers, and knowledge of math content by math teachers.

Next, we estimate the contribution of teacher content knowledge to student learning at the student-by-subject level.[1] We estimate the following regression:

$$Stu\_EndScore_{isj} = \beta Tea\_Knowledge_{sj} + \gamma Stu\_BaseScore_{isj} + \psi_{isj} + \delta_s + \mu_{isj} \tag{2}$$

* Beg: University of Delaware and J-PAL. Fitzpatrick: The Ohio State University. Kerwin: University of Washington, IZA, and J-PAL. Lucas: University of Delaware, J-PAL, and NBER. Rahman: University of Oxford and BIGD, BRAC University. All study activities were approved by the following Institutional Review Boards: IFMR, University of Minnesota, University of Delaware, and University of Massachusetts Boston. This evaluation would not be possible without our partners at Transform Schools, People for Action, and the School and Mass Education Department of Odisha, including Pankaj Sharma, Bindiya Nagpal, Christine Oliver, and Kartik Sahu. We also thank the JPAL-South Asia field operations staff, and all of our respondents for their time, cooperation, and insight. Soumya Das provided excellent research assistance. Funding for this project was provided by the UK Foreign, Commonwealth & Development Office, awarded through J-PAL's Post-Primary Education Initiative, and the Kusuma Trust UK. AEA RCT #AEARCTR-0004138. None of the authors have any potential conflicts of interest.

[1] We limit our estimation to students with baseline and endline test scores for both English and math. A limited number of students only took one test.

where $Stu\_EndScore_{isj}$ is the endline score for student i in subject s taught by subject teacher j, $Stu\_BaseScore_{isj}$ is the student's subject-specific baseline test score, ψ_{isj} are student fixed effects, and δ_s are subject fixed effects.[2] We use the same continuous measure of *Tea_Knowledge* as in equation (1). We cluster our standard errors at the level of a teacher, which is the level at which our regressor of interest varies.[3]

Following the teacher value-added literature, we use student fixed effects to control for student-level time-invariant omitted variables, including the sorting of students to teachers or schools (Chetty, Friedman, and Rockoff, 2014).[4]

II. Data

We use data from a randomized evaluation of *Utkarsh*, a secondary-school remedial-education program in Odisha, India in Class 9 (ninth grade) in 300 schools during the 2019-2020 school year (Beg et al., 2024, 2026). Utkarsh sets aside entire school days to intensely focus on different levels of the curriculum from foundational learning through grade level.[5] The evaluation found that replacing part of the standard curriculum with targeted remedial instruction improved student learning by 0.11 standard deviations (roughly 60% of a year of *status quo* learning). These gains did not crowd out grade-level mastery, and occurred even when teachers were given the flexibility to deviate from the prescribed lesson plans. Our sample includes both a control group that continued under the *status quo* and schools that received Utkarsh training and materials. We use all schools to maximize statistical power, breaking out our analyses separately by treatment status in some cases.

In this paper, we focus on surveys and English and math test scores of teachers and students. [6] We measure teacher content knowledge using a previously validated grading-based method: teachers grade a mock student assignment in mathematics or English (Bold et al., 2017). A teacher's content knowledge score is the fraction of the questions on the student assignment that they grade correctly. A teacher is deemed to meet a minimum level of content knowledge if they score 80% or higher (Bold et al., 2017).[7] This test is at the 4th grade level, a lower bound for the content that teachers should know about their subjects. Further, in our setting, the average Class 9 student scored at about 4th grade proficiency at baseline, making teachers' 4th grade content knowledge particularly salient.

We assessed the English and math knowledge of the 899 teachers present during the second wave of our endline survey (226 English teachers, 302 math teachers, 3 teachers of math and English, and 368 teachers of other subjects). The average teacher is 41 years old and 54% of them are male. Teachers are highly educated (52% have a master's degree or higher) and quite experienced (on average 16 years).

III. Results

A. Levels of Teacher Content Knowledge

In Figure 1, we present the distribution of teacher content knowledge by subject, separately by subject taught. Ninety percent of Class 9 math teachers and only 5% of Class 9 English teachers

[2] We control for subject fixed effects (δ_s) to control for potential bias because our teachers score systematically higher on math than on English and our students had more growth in math than in English. In some specifications we include school fixed effects instead of student fixed effects.

[3] Two-way clustering by both teacher and student or clustering at the school level yielded similar results. Two-way clustering did not converge in some cases.

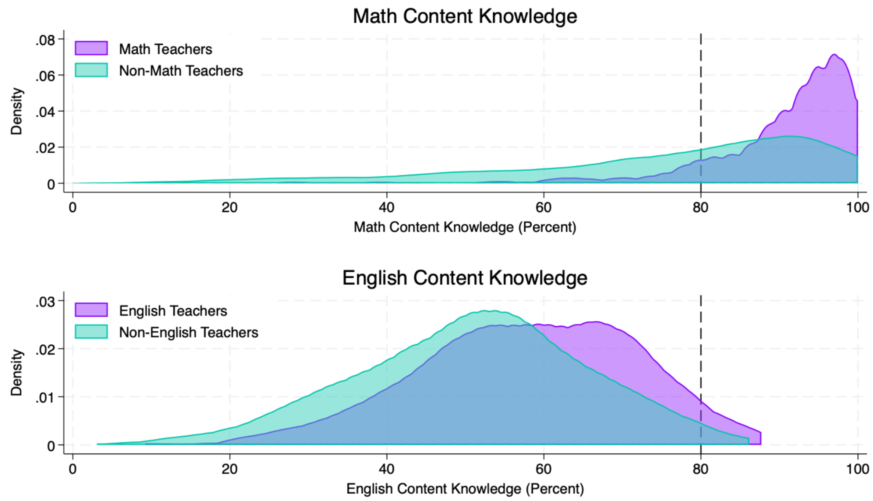
[4] We can include these fixed effects because we have multiple subject-specific teachers and test scores for each student.

[5] At the school level, We tested two versions of the Utkarsh treatment: standard (strict prescription of a daily implementation schedule) and flexible (allowing teacher discretion over parts of the schedule). For this paper we combine the two versions.

[6] We also invigilated student tests in Science and Odia. In this paper we focus only on math and English as those are the subjects covered by the validated teacher content knowledge tests. Student test scores are based on exams that included both grade level and remedial material. Scores are calculated using item response theory and standardized relative to the baseline mean and standard deviation.

[7] The Bold et al. (2017) paper formally tests "language of instruction," for primary school teachers. While the language of instruction in Odisha is Odia, we test teachers on content knowledge in English, as secondary school teachers are subject teachers.

FIGURE 1. DISTRIBUTION OF TEACHER CONTENT KNOWLEDGE SCORES



Note: Above is the kernel density of content knowledge scores. The top graph shows the results on the math content knowledge test disaggregated by math and non-math teachers. The bottom graph shows the results on the English content knowledge test disaggregated by English and non-English teachers. The vertical line corresponds to 80%, the cut-off score for whether a teacher meets the minimum content knowledge threshold from Bold et al. (2017). $N = 899$. Math content knowledge: non-math teachers mean=75.5, SD=21.2; math teachers mean=90.7, SD=10.9. English content knowledge: non-English teachers: mean=50.8, SD=14.8; English teachers: mean=58.0, SD=13.9.

meet the 80% minimum content knowledge standard for 4th grade. Content knowledge is higher for the teacher's subject of specialization, more so for math. Relative to 4th grade teachers in the seven African countries studied in Bold et al. (2017), Indian Class 9 teachers score lower on the language test but higher on the math.[8]

B. *Correlates of Teacher Content Knowledge*

Table 1 presents correlates of teacher content knowledge, estimating Equation (1) separately for English (column 1) and math teachers (column 2) on demographic variables. There are few statistically significant correlates of teacher content knowledge in our sample: education and years of experience for English teachers, and gender and years of experience for math teachers.[9] The overall explanatory power is also low for both.

C. *Regressions of Student Test Scores on Teacher Content Knowledge*

Table 2 shows the relationship between student test scores and teacher content knowledge based on Equation (2). Column 1 controls only for subject fixed effects (i.e., English and math), column 2 adds school fixed effects, and column 3 replaces school fixed effects with student fixed effects. We show results for three samples: the full sample (Panel A), control schools only (Panel B), and treatment schools only (Panel C).

In the full sample (Panel A), teacher content knowledge increases student test scores once we control for either school or student fixed effects (columns 2 and 3), which are our preferred specifications as they isolate content knowledge from other confounders. In the control schools (Panel B), there is

[8]The countries in Bold et al. are Kenya, Mozambique, Nigeria, Senegal, Tanzania, Togo, and Uganda. For language, 66% of teachers in the African set meet the minimum level of content knowledge (ranging from 26% in Nigeria to 94% in Kenya), exceeding our sample by on average by 61 percentage points. For math, 68% of teachers in the African sample meet the minimum content knowledge standard (ranging from 49% in Togo to 93% in Kenya), 22 percentage points lower on average than our sample.

[9]These patterns are quantitatively similar when we use only the control group, but noisier due to the smaller sample.

TABLE 1—CORRELATES OF TEACHER CONTENT KNOWLEDGE

	English Score (PCT) (1)	Math Score (PCT) (2)
Male	-0.883	3.547
	(2.569)	(1.687)
Age	-1.372	-1.742
	(1.053)	(1.452)
Age-Sq	0.012	0.012
	(0.012)	(0.017)
Education: Bachelors or Lower	-3.278	0.671
	(1.967)	(1.394)
Years Experience	1.283	1.231
	(0.526)	(0.736)
Years Experience-Sq	-0.023	-0.021
	(0.011)	(0.019)
Observations	229	305
R^2	0.050	0.079
Mean, Dep Var	57.985	90.748

Notes: Outcome variable: teacher's content knowledge score, measured 0 to 100. Treatment status included as additional control variables. Heteroskedasticity-robust standard errors appear in parentheses. Column 1: outcome is English content knowledge score, sample is English teachers. Column 2: outcome is math content knowledge score, sample is math teachers.

TABLE 2—IMPACT OF TEACHER CONTENT KNOWLEDGE ON STUDENT TEST SCORES

	Student Test Scores (SD)		
Panel A: All Schools	(1)	(2)	(3)
Content Knowledge	0.008	0.088	0.087
	(0.057)	(0.035)	(0.043)
	[0.893]	[0.011]	[0.045]
# Students	9,316	9,316	9,316
R^2	0.747	0.770	0.934
Panel B: Control Schools			
Content Knowledge	-0.057	0.055	0.023
	(0.099)	(0.055)	(0.061)
	[0.568]	[0.323]	[0.705]
# Students	3,232	3,232	3,232
R^2	0.740	0.773	0.939
Panel C: Treatment Schools			
Content Knowledge	0.059	0.117	0.137
	(0.064)	(0.043)	(0.058)
	[0.357]	[0.007]	[0.019]
# Students	6,084	6,084	6,084
R^2	0.749	0.766	0.930
School FE	No	Yes	No
Student FE	No	No	Yes

Notes: Content knowledge measured as proportion correct from 0 to 1. Student test scores are standardized relative to the baseline mean and standard deviation. Heteroskedasticity-robust standard errors, clustered by teacher, in parentheses with associated p-values in square brackets. All regressions include baseline student test scores and subject fixed effects. Column 1 in Panel A also controls for treatment status.

a robust null effect of teacher knowledge on student test scores. In treatment schools only (Panel C), in contrast, there is a significant positive effect: increasing competence by 100 percentage points increases student test scores by 0.14 SD (column 3). In our sample, no teachers scored 0; a move from the minimum (0.23) to the maximum knowledge score (1.00) increases test scores by 0.11 SD, and a 1-SD increase in teacher knowledge (0.20) raises student test scores by 0.03 SD.

To interpret these results, we first note that the Utkarsh program has almost no effect on teacher content knowledge.[10] Our results suggest that the treatment may have changed how teacher knowledge affects student test scores. Under the *status quo*, teacher knowledge does not matter for student learning. This may be due to a focus on rote memorization, where teachers lecture students on fixed content with minimal feedback from students. Teachers operating in this system may have similarly memorized the content by rote, rendering their understanding of it irrelevant. In contrast, the remedial focus of Utkarsh relies more heavily on teacher knowledge of the foundational material tested in the teacher exam, and teachers have more opportunity to impart that knowledge to students via more meaningful student-teacher interactions.

IV. Discussion and Conclusion

We measure secondary school teacher content knowledge in Odisha, India. Despite most teachers holding a Master's degree, the majority of Class 9 English teachers in our sample do not meet a fourth-grade content knowledge standard. A smaller proportion of math teachers also fail to reach this content threshold. Teachers' knowledge about their subjects is at most weakly predicted by their demographics.

Teacher content knowledge is unrelated to student learning under the *status quo*. In contrast, in schools where teachers received training and materials for a remedial education intervention, higher teacher content knowledge results in more student learning.

These results have two important policy implications. First, some of the heterogeneity across settings in the degree of success of differentiated instruction programs could be due to variation in teacher content knowledge—grade-level subject teachers may not be proficient in remedial content themselves. Second, to increase the effectiveness of remedial programs, teachers could need additional training in both pedagogy and content. Future work should explore both of these possibilities.

REFERENCES

Beg, Sabrin A., Anne E. Fitzpatrick, Jason T. Kerwin, Adrienne M. Lucas, and Khandker Wahedur Rahman. 2024. "When Given Discretion Teachers Did Not Shirk: Evidence from Remedial Education in Secondary Schools." National Bureau of Economic Research Working Paper 33242.

Beg, Sabrin, Anne Fitzpatrick, Jason Kerwin, Adrienne Lucas, and Khandker Wahedur Rahman. 2026. "Data for: "Teacher Content Knowledge in Indian Secondary Schools and Its Relationship with Student Learning"." Distributed by the J-PAL Dataverse.

Bold, Tessa, Deon Filmer, Gayle Martin, Ezequiel Molina, Brian Stacy, Christophe Rockmore, Jakob Svensson, and Waly Wane. 2017. "Enrollment without Learning: Teacher Effort, Knowledge, and Skill in Primary Schools in Africa." *Journal of Economic Perspectives*, 31(4): 185–204.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review*, 104(9): 2633–2679.

World Bank. 2018. *World Development Report 2018: Learning to Realize Education's Promise.* Washington, DC:World Bank.

[10]See Appendix Table B20 of Beg et al. (2024). The only statistically significant effect is that Flexible Utkarsh (the second treatment group) increases math content knowledge by 3 percentage points relative to a control-group mean of 89%. There are no effects for English or for the other treatment arm.