IEA POLICY BRIEF

Researching education, improving learning

NUMBER 16 OCTOBER 2017

The role of teacher quality in fourth-grade mathematics instruction: evidence from TIMSS 2015

BY **NATHAN BURROUGHS** & **AMITA CHUDGAR** (MICHIGAN STATE UNIVERSITY)

SUMMARY

Does teacher quality differ within and between countries, and how are measures of teacher quality related to instructional alignment and instructional time in mathematics? Fourth-grade classroom data from the IEA's Trends in Mathematics and Science Study (TIMSS) 2015 revealed that although measures of teacher quality were only weakly related to one another, countries with more variation along one dimension (for example, experience) also have high variation along other dimensions (such as education, or readiness to teach math topics). Measures of teacher quality were not strong or consistent predictors of instructional alignment or time, suggesting that primary school teachers' preparation to teach mathematics may have limited influence on classroom opportunity to learn. Crucially, in many countries, disadvantaged students have (by some measures) higher quality teachers. Teacher collaboration and school expectations of instructional behavior merit further research, and the influence of principals and other teachers may have strong impact on educational outcomes. Policymakers should focus their efforts on improving their systems of teacher preparation.

THE CHALLENGE: IMPROVING TEACHER QUALITY

In their efforts to improve educational systems, school leaders and policymakers have long sought to influence what happens in the instructional core of education: the interaction between teachers, students, and instructional content. Educational leaders have to a varying degree attempted to influence instructional effectiveness through the use of standards, assessments, and prescribed curricula.

However, in practice, teachers have wide discretion in their classrooms. Consequently, improving teacher effectiveness remains one of the most tantalizing challenges confronting educational policymakers around the globe. Despite considerable research attesting to the importance of teachers to student achievement (see for example Chetty et al., 2011; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004), there is still no clear consensus about what constitutes a good teacher or how to train one. Further, most existing research is restricted to specific countries (in particular the USA), raising questions about applicability when other countries may have different contexts[1]. Policymakers thus lack clear guidance, while policies have remained as muddled and inconsistent as the research.

1 A notable exception is the IEA's TEDS-M (Teacher Education and Development Study in Mathematics), which examined teacher preparation in 17 countries (Ingvarson et al., 2013).

CONTENTS

TRADITIONAL MEASURES OF TEACHER QUALITY: EDUCATION AND EXPERIENCE

Traditionally, policymakers have relied on education and experience as the principal measures of teacher quality. In the USA, for example, many states adopted more generous salary policies and stricter certification requirements in an effort to improve the human capital of their teacher labor force, and thereby improve student outcomes.

However, researchers have generally failed to find a clear link between these teacher characteristics and student achievement in standardized assessment. International comparative studies have failed to identify a consistent relationship between student achievement and teacher education (measured by advanced degrees, subject specialization, and certification) or teacher experience (see for example Luschei, & Chudgar, 2011). Gustafsson and Nilson (2016) found that teacher education was related to student outcomes in multiple countries, while in the USA, Ladd and Sorenson (2016) demonstrated teacher efficacy was linked to their length of job service. Nevertheless, the effect of teacher education and experience on student achievement remains unclear.

A FOCUS ON WHAT TEACHERS DO? INSTRUCTIONAL TIME AND INSTRUCTIONAL CONTENT

Are teacher education and experience appropriate measures for such a broad concept as "teacher quality," or are there better alternatives, especially given the great expense in attempting to enhance these qualities in teachers? Although teacher content knowledge and instructional effectiveness are difficult to measure directly, they have been associated with professional development (Desimone, Porter, Garet, Yoon, & Birman, 2002) and teacher preparedness (Blomeke, Olsen, & Suhl, 2016), which could serve as indirect measures that are more easily collected.

Ultimately, policymakers are probably less interested in teachers' attributes than what they do in the classroom. Two factors are crucially important: instructional content and instructional time. There is considerable cross-national evidence that exposure to rich content has a major impact on student learning (Schmidt, Burroughs, Zoido, & Houang, 2015). Students also do better if more classroom time is devoted to a subject (Lavy, 2015). Recognizing that instructional time and content have the potential to influence student outcomes, leaders in many educational systems have mandated the amount of time devoted to content-specific instruction and attempted to strengthen instructional content through standards-based reforms.

Along with a focus on achievement, teacher characteristics have also been viewed as an important means of addressing educational inequality. In the USA, for example, the distribution of experienced and better educated teachers appears to exacerbate educational inequality (Goldhaber, Lavery, & Theobald, 2015), but other countries do a better job of assigning their strongest teachers to their more disadvantaged students (Akiba, LeTendre, & Scribner, 2007). Teacher behaviors like instructional content and instructional time may also contribute to educational inequality.

HOW MUCH LEVERAGE DOES POLICY HAVE TO INCREASE OPPORTUNITY TO LEARN?

Although the direct influence of teacher experience and education on student learning is tenuous, more experienced and better prepared teachers may engage in more effective teaching practices; in short there might be an *indirect* link between traditional measures of teacher quality and student outcomes. In addition, it is also possible that along with these input-based effects (which are difficult to change), instructional effectiveness might also be improved by professional development and standards-based reform, factors more malleable to policy changes (Figure 1).

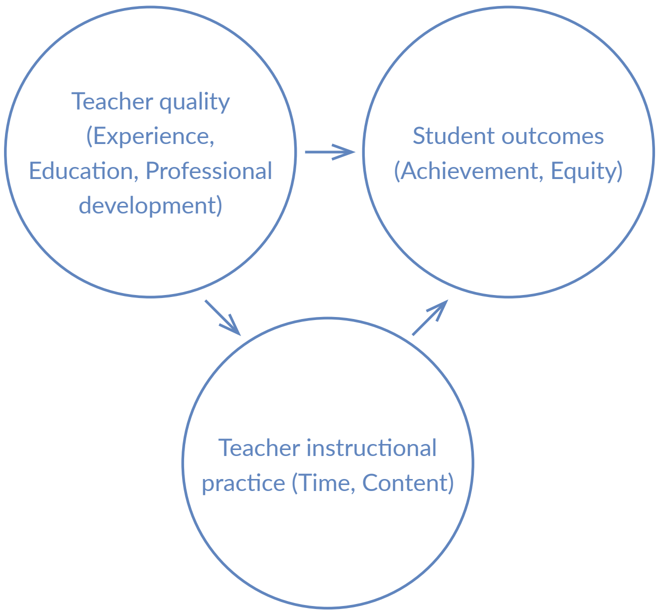

Figure 1: Model of the interrelationship of teacher quality, teacher instructional practice, and student outcomes

The focus of this study is on fourth-grade classrooms, where the interrelationship of teacher quality, instructional practice, and student outcomes may be stronger. Classroom autonomy is likely to be far greater, and hence teacher knowledge and predispositions more powerful, in early grades. Primary classrooms are usually characterized by whole-classroom instruction in which one teacher teaches every subject, giving them discretion over instructional content and instructional time.

DATA AND KEY MEASURES

We explored the distribution of measures of teacher quality and their relationship to time on task and alignment with national standards using data for fourth-grade mathematics students from the 2015 TIMSS study (https://timssandpirls.bc.edu/timss2015/; http://www.iea.nl/data). Our key measures of teacher quality were: pre-service teacher preparation in mathematics, self-reported readiness to teach mathematics topics, experience teaching, receipt of professional development in mathematics, instructional time on mathematics, and percentage alignment to national curriculum standards. Detailed information about how we defined each of these measures is available as an Appendix (see www.iea.nl/sites/default/files/publications/Electronic_versions/PolicyBrief16App.pdf). When interpreting these results, note that the teachers are not a random sample within each country, but are the teachers of a random sample of students.

QUESTION 1: WHAT IS THE DISTRIBUTION OF TEACHER QUALITY WITHIN AND ACROSS COUNTRIES?

We established six measures of teacher quality for 48 educational systems (hereafter referred to as countries) participating in the TIMSS (Table 1); for each country, we calculated averages, standard deviations, and socioeconomic status (SES) inequality. Country averages revealed significant national differences in the teachers' levels of mathematics education; the Netherlands scored 4.65, indicating that most teachers have an advanced degree and specialized in mathematics, while the scores of Italy (1.46) and Morocco (1.61) indicate that many primary school math teachers in these countries have little more than a high school diploma. The average fourth-grade math teacher had 17 years of teaching experience; a block of Eastern European countries (Lithuania, Bulgaria, Russia, and Hungary) reported higher average levels of teacher experience. Teachers' reported preparation to teach mathematics topics varied less across countries, with most students' teachers indicating that they felt prepared to teach most topics (the international average was 2.81, range 2.34 to 2.99). By contrast, there was a massive variation in receipt of professional development, from Cyprus' 0.8 to its neighbor Turkey's 0.06.

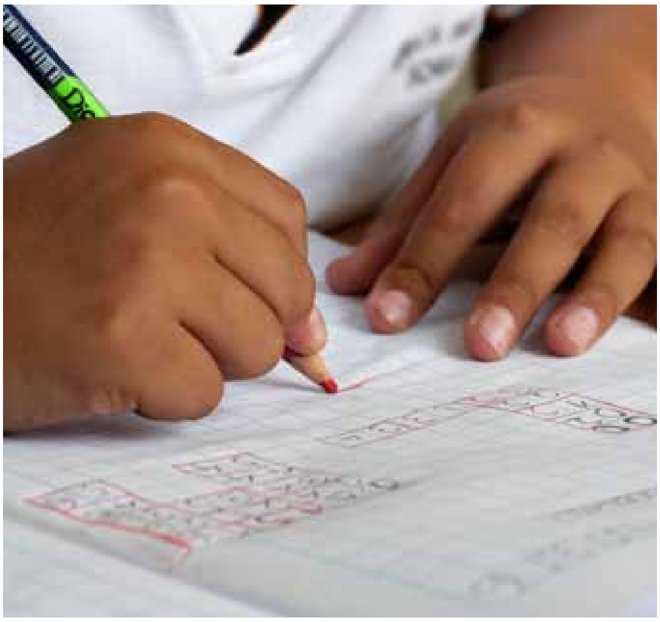

Table 1: Six measures of teacher quality

		Preparation (level of education)	Experience	Readiness	Professional development	Instructional Time	Alignment (with national curriculum)
Country means	Avg	3.85	16.80	2.81	0.42	253	0.53
	Min	1.46	9.05	2.34	0.06	157	0.40
	Max	4.65	27.65	2.99	0.80	478	0.67
Standard deviation	Avg	0.73	7.69	0.24	0.28	46	0.12
	Min	0.06	0.95	0.02	0.04	7	0.02
	Max	5.01	38.70	1.84	1.72	314	0.55
SES gaps	Avg	0.05	0.60	0.04	-0.01	-8	0.00
	Min	-0.62	-8.90	-0.15	-0.19	-70	-0.08
	Max	1.69	8.20	0.33	0.22	25	0.08

Notes: Avg = average, Min = minimum, Max = maximum, SES = socioeconomic status. Instructional time is coded in minutes; experience is coded in years; readiness is derived from responses to 17 questions on a three-point scale, alignment is coded as percentage of teachers aligned with national curriculum standards; preparation on a five-point scale; professional development (PD) as mean receipt of PD on a 1/0 scale. See Appendix (www.iea.nl/sites/default/files/publications/Electronic_versions/PolicyBrief16App.pdf) for a full explanation of the measures and all the country level data used in the calculations.

Portuguese teachers spent the most time teaching mathematics (averaging 478 minutes of mathematics instruction per week), and Korea's teachers the least (157 minutes per week); the average was 253 minutes, or slightly more than 4 hours per week. This suggests that the typical fourth-grade student receives less than an hour of math instruction per school day, with Portuguese students receiving over an hour and a half per day and Korean students about half an hour. Most countries reported expectations for time allotted to mathematics in the TIMSS national context survey, and these expectations are correlated with teacher-reported time devoted to mathematics (Pearson's correlation of 0.57).

There is also modest average alignment with a country's national standards. Teachers of Korean (0.67) and US (0.66) students were most aligned with national standards, with Portugal and Finland least aligned (0.4). It is notable that the USA, despite not having a formal national curriculum, still demonstrated relatively high alignment with suggested topic coverage at fourth grade.

Although all of these metrics can be interpreted as measures of teacher quality, they are really quite distinct. Rather than being strongly associated, the relationship between the various measures of teacher quality are fairly weak, and in several cases negative. For example, the strongest relationship is between teacher experience and receipt of professional development, but this is negative (–0.47). This may mean that countries with less prepared fourth-grade teachers devote more time to training teachers once they are in the classroom. Using national averages, there is little support for the hypothesis that more experienced and educated teachers devote more time to teaching mathematics or are better aligned with national standards; in fact, those correlations are weakly negative. However, these correlations use no statistical controls and it is likely that most of the variation is within country rather than between countries.

We found substantial within-country variation in measures of teacher human capital and significant standard deviation in teacher quality metrics (Table 1). In contrast to country means, countries with the largest variation in one dimension of teacher human capital also showed substantial variation in other dimensions. The measures are all highly correlated (0.83–0.96), and factor analysis produced a single dimension accounting for 93% of the variance. Countries with larger between-classroom variation in teacher quality also recorded bigger differences in the average SES of classrooms.

This implies that inequality in teacher human capital systematically tracks student poverty, since considerable research (especially in the USA) indicates that the most disadvantaged students are more likely to be paired with the lowest quality teachers. However, this speculation would be incorrect. In calculating the difference in teacher quality between the highest SES and lowest SES quartiles of classrooms (in other words, between the richest 25% of classrooms and the poorest 25% of classrooms), the results indicate that many countries assign more experienced or better educated teachers

Table 2: Statistically significant relationships between teacher quality and teacher instruction

Measures of teacher quality	Measures of teacher instruction			
	Instructional time		Alignment	
	+	–	+	–
Preparation	England Italy Turkey	N. Ireland Singapore	Finland	Australia Oman
Experience	Italy	England Spain Turkey	Cyprus Lithuania Spain	
Readiness	Ireland Oman Qatar Singapore UAE	Cyprus	Canada Cyprus Iran New Zealand	Lithuania
Professional development	Spain USA	Lithuania Portugal	Oman Qatar	Canada
Instructional time			Bahrain Finland Italy Qatar Turkey	

Notes: **+** = statistically significant positive relationship with instructional time or alignment.
- = statistically significant negative relationship with instructional time or alignment.
UAE = United Arab Emirates.
Only countries with a statistically significant relationship are reported. The analysis includes a sample of 28 countries, with 13 eliminated due to a small sample size. The average adjusted r-square was 0.08. (Complete country level data are available in the Appendix at www.iea.nl/sites/default/files/publications/Electronic_versions/PolicyBrief16App.pdf.)

to their most disadvantaged rather than their most advantaged students.

Differences in teacher quality between richer and poorer classrooms are weakly related across different metrics of teacher quality, and also weakly correlated to overall within-country variation. To put it more simply, the consistently large variation in teacher quality found in some countries is not because weaker teachers are assigned to classrooms with poorer students. In fact, even in the USA, the TIMSS data suggests that the lowest quartile of students have teachers that are more experienced, spend more time on math, receive more professional development, and are better aligned with national standards (but have less education and confidence to teach math). Of course, these findings are sharply at odds with much of the existing literature on teacher quality in the USA, and further study is thus warranted.

QUESTION 2: WHAT IS THE RELATIONSHIP BETWEEN TEACHER QUALITY AND INSTRUCTIONAL TIME?

We next explored whether fourth-grade mathematics teachers with more preparation and confidence in mathematics spent more of their class time on math. This question has important policy implications, because if traditional measures of teacher quality (education and experience), teacher content knowledge (readiness), or receipt of professional development in mathematics are associated with more time on task, then policymakers would have important information about crafting useful interventions. However, if these characteristics are not systematically related to instructional time, then policymakers and school leaders will have to shift to other strategies for intervening in the instructional core rather than relying on increasing teacher human capital.

We ran a series of regressions[2] predicting classroom time on math using five predictors: education, experience, professional development, self-reported readiness to teach math, and a control variable, student socioeconomic status, as measured by the reported number of books in the home. Unfortunately, 13 countries had to be eliminated from the analysis because there were fewer than 10 observations per parameter (teachers did not respond to all questions): this (1) reduces the power of the analysis, and (2) raises the possibility of selection bias.

Echoing the findings of Luschei and Chudgar (2011), we found no consistent relationship between teacher characteristics and teacher time on mathematics (Table 2). More time on mathematics was statistically significantly associated with teacher education in four systems, experience in one system, readiness in five systems, and professional development in two systems. In several countries, the association with instructional time was statistically significant and negative; higher teacher human capital was associated with less time spent teaching math.

The substantive implications of these results are twofold. First, there is no clear, consistent pattern between teacher characteristics and instructional time on task across countries. For example, despite their similarities, what holds for Canada most decidedly does not hold for the USA. Our findings show that the institutional context is quite distinct. Second, as a general rule, teacher inputs and broad interventions such as professional development do not by themselves lead to substantial increases in teacher instructional time for mathematics.

QUESTION 3: WHAT IS THE RELATIONSHIP BETWEEN TEACHER QUALITY AND INSTRUCTIONAL ALIGNMENT

National curriculum standards are an important mechanism employed by most countries to guide teachers in conveying age-appropriate content. In principal, standards are organized so that mathematics topics flow from more basic foundational material in earlier grades to more sophisticated mathematical concepts in later grades. Under a well-designed set of standards, it is crucial that teachers fully instruct their students in the assigned grade-level content, so that teachers in later grades can focus on the topics assigned to those grades rather than devoting time to teaching material that should have been already covered.

As indicated earlier, there are considerable discrepancies between the content that teachers are expected to teach and what they actually report teaching. It may be that teachers at the primary level do not feel comfortable with mathematics, have limited familiarity with the standards, or lack the knowledge to adhere to them. It is also plausible that more experienced teachers have greater knowledge of mathematics in general and national standards in particular. If human capital contributes to instructional alignment, then by investing in teacher characteristics (via retention strategies, preparation to teach, or opportunities for professional development) policymakers could reduce the difference between the

2 Because one of the purposes of this paper is to explore how these within-country relationships differ across countries, we ran regressions for each country, rather than one combined multi-level model.

intended and implemented curriculum. However, the analysis does not support this hypothesis.

We found a weak association between teacher quality metrics and curricular alignment (Table 2). Of the 28 TIMSS systems for which there was sufficient sample to make an estimate, greater alignment was statistically significantly related to higher education in one, experience in four, readiness in five, professional development in two, and time on mathematics in six countries. Again, in several instances there were negative statistically significant relationships, and the predictive power of the model ranges from 1% to 15% of the total variance (averaging 0.08). As with instructional time on mathematics, the degree to which fourth-grade mathematics teachers' coverage of mathematics topics matches what is called for in national standards appears to be largely independent of their knowledge, experience, and professional support.

CAVEATS

There are a number of limitations to this study, and we urge caution in interpreting these results. First, all of the teacher measures are self-reported and are based on question items with a fairly narrow range, and with fairly broad categories. This may lead to lack of precision in the estimates and measurement error. Second, the study is not based on a random sample of teachers, but the teachers of a random sample of students. Third, we discarded some of the original sample because of low response rates and, if failure to answer is related to teacher knowledge or preparation, this could bias the estimates. Fourth, the analytical approach is cross-sectional and correlative rather than longitudinal and causal, and does not control for a number of other factors, especially school factors. The TIMSS study samples one or two classrooms in a school, and it is thus difficult to separate out the influence of school contextual factors from specific classroom factors.

REFERENCES

Akiba, M., LeTendre, G., & Scribner, J. (2007) Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher*, *36*(7), 369-387.

Blomeke, S., Olsen, R., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In T. Nilson, & J.-E. Gustafsson (Eds), *Teacher Quality, Instructional Quality and Student Outcomes* (pp. 21-50). IEA Research for Education, Volume 2. Cham: Springer Open.

Chetty, R., Friedman, J., Hilger, N., Saez, E., Schanzenbach, D., & Yagan, D. (2011). How does the kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, *126*(4), 1593-1660.

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

CONCLUSIONS FOR POLICYMAKERS

- There is significant variation in teacher quality within countries. Although teacher quality and instructional metrics are not strongly related to one another, *variation* in these measures are: countries with large variation in teacher experience also tend to exhibit large variation in teacher education, readiness, and instructional time. This could partly be a product of the system of teacher preparation. For example, educational systems where teacher education is more tightly regulated might be reasonably expected to exhibit lower variation. Efforts to improve teacher effectiveness should attend to its multidimensionality.
- Disadvantaged students are not necessarily assigned lower quality teachers. Conventional wisdom is that low-SES students inevitably receive weaker instruction, delivered by less experienced, less prepared teachers. However, the TIMSS data suggests that the average teacher human capital of low-income students (bottom 25%) is often higher than for higher-resourced students, even, by some measures, in the USA. In other words, the systematically lower teacher quality found in many studies of disadvantaged students is not inevitable, and is amenable to policy change.
- Teacher characteristics are weakly associated to teacher behavior related to opportunity to learn. We failed to find a strong link between teacher inputs (or even professional development) and time on math or alignment to national standards. This may be because self-reported readiness to teach and teacher education are weak proxies for teacher content knowledge or instructional quality. Or it may be that school culture is far more important that easily observable teacher characteristics in determining teacher effectiveness. Whatever the direct effects that experience or education may have on student learning, instructional alignment and instructional time are independent factors, which means that policymakers and researchers will have to develop interventions that are designed explicitly to increase alignment and instructional time. Greater human capital is unlikely to accomplish the task.
- One element that warrants greater attention is teacher collaboration and school expectations of instructional behavior, both of which could have a major impact on what teachers do in the classroom. Our analysis largely treats classrooms as isolated units existing within a national educational framework, but it is very likely that principals and other teachers have greater scope to influence a teacher's behavior than the mandates of education ministries. Policymakers may have more success in influencing the instructional core by focusing their efforts on the schools as a whole, along with improving their systems of teacher preparation.

Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Education Researcher, 44*(5), 293-307.

Gustafsson, J.-E. & Nilson, T. (2016). The impact of school climate and teacher quality on mathematics achivement: a difference-in-differences approach. In T. Nilson, & J.-E. Gustafsson (Eds), *Teacher Quality, Instructional Quality and Student Outcomes* (pp. 81-95). IEA Research for Education, Volume 2. Cham: Springer Open.

Ingvarson, L., Schwille, J., Tatto, M.T., Rowley, G., Peck, R., & Senk, S.L. (2013). *An analysis of teacher education context, structure, and quality-assurance arrangements in TEDS-M countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam: IEA.

Ladd, H.F., & Sorenson, L.C. (2016). Returns to teacher experience: student achievement and motivation in middle school. *Education Finance and Policy*, *12*, 241-279. doi:10.1162/EDFP_a_00194.

Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *The Economic Journal*, *125*, 397-424.

Luschei, T., & Chudgar, A. (2011). Teachers, student achievement, and national income: a cross-national examination of relationships and interactions. *Prospects*, *41*, 507-533.

Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417-458.

Rockoff, J. (2004). The impact of individual teachers on student achievement: evidence from panel data. *The American Economic Review*, *94*(2), 247-252.

Schmidt, W., Burroughs, N., Zoido, P., & Houang, R. (2015). The role of schooling in perpetuating educational inequality: an international perspective. *Education Researcher*, *44*(4), 371-386.

IEA POLICY BRIEF

ABOUT THE IEA

The International Association for the Evaluation of Educational Achievement, known as the IEA, is an independent, international consortium of national research institutions and governmental agencies, with headquarters in Amsterdam. Its primary purpose is to conduct large-scale comparative studies of educational achievement with the the aim of gaining more in-depth understanding of the effects of policies and practices within and across systems of education.

ISSN: 2215-0196
Photo credits:
CC0 Public Domain
Copies of this publication can be obtained from:
IEA Amsterdam
Keizersgracht 311
1016 EE Amsterdam
The Netherlands
By email: secretariat@iea.nl
Website: www.iea.nl

Follow us @iea_education

Editor of the policy brief

David Rutkowski
Centre for Educational Measurement at the University of Oslo (CEMO)

Please cite this publication as:
Burroughs, N., & Chudgar, A. (2017, October). The role of teacher quality in fourth-grade mathematics instruction: evidence from TIMSS 2015 (Policy Brief No. 16). Amsterdam, The Netherlands: IEA.