ISSUE 7 JULY 2017

CAMBRIDGE Mathematics ESPRESSO

RESEARCH, FILTERED BY CAMBRIDGE MATHEMATICS

TALKING POINT:

HOW CAN MATHEMATICS TEACHING BE MEASURED?

IN SUMMARY

- It is hard to measure good maths teaching
- Triangulation of student attainment, student evaluation and lesson observation provides the most stable measure of maths teaching
- Lesson observation alone is unreliable, but can be improved by using multiple observers
- Several possible instruments have been developed to help measure the effectiveness of maths teaching; choice of these should be based on an identified purpose of measurement
- Decisions about teachers' employment are currently often based on inadequate measures or proxies
- Factors other than teacher effectiveness such as schools affluence and resources have effects on pupil outcomes

'it may be difficult to disentangle the effect of teacher quality and student characteristics on teachers' value-added scores'
Hill et al

'Using Ofsted's categories, if a lesson is judged 'Outstanding' by one observer, the probability that a second observer would give a different judgement is between 51% and 78%.'
Prof Rob Coe

<50%

accuracy of maths lesson observation judgements by well-trained observers using a nationally recognised instrument (data from Strong *et al*, 2011)

50%

chance of correctly guessing the result of a coin toss

67%

chance of getting the prize if you switch choices in the Monty Hall problem

1

There are many ways to measure mathematics teaching, including lesson observations (with or without an instrument), student attainment (including value-added models) and student evaluation. The most consistent and stable measures of good mathematics teaching combine all three[1] by careful triangulation[2]. The least reliable evidence is provided by looking at lesson plans, student work, marking, and self-reports by teachers[2]. Some studies define good maths teaching as that which produces the best student outcomes, although this is dependent on the quality and purpose of assessment. Although perhaps seen as the most accurate measure, value-added models are still likely to be biased by prior teaching effects, distribution of students, and measurement error[3, 2]. It is difficult to disentangle the effects of school, teacher and year group[4].In all of these teacher-level variance is much higher than year group, school or local area-level variance[5] – in short, good mathematics teaching seems to make a real difference[1], but is hard to measure.

IMPLICATIONS: Good maths teaching makes a difference but is hard to measure

Lesson artefacts and self-reports by teachers are not effective measures of maths teaching

The most stable measures of effective maths teaching use a triangulation of observations, attainment and student evaluation

CAMBRIDGE UNIVERSITY PRESS
UNIVERSITY OF CAMBRIDGE Faculty of Mathematics
UNIVERSITY OF CAMBRIDGE Faculty of Education
Cambridge Assessment

2

Observing teachers is a process which usually involves judging against a set of teacher standards; creating these is an inherently political process and the final standards are a compromise of research, opinion and intuition[5]. Observing lessons is often unreliable[6] and should be used with caution and not as the basis for important decisions; however training observers and using protocols improves reliability against the defined criteria[2]. If using lesson observation as a measure of teacher effectiveness, using a second observer rather than the same observer again increases reliability[1, 2]. Narrow measures of effective teaching like specific checklists tend to incentivise specific behaviours in teachers and make it harder to identify other important outcomes[1, 2].

IMPLICATIONS: Lesson observation is prone to particular bias in terms of the focus of standards by which it is measured and is generally unreliable

Reliability can be increased by using multiple observers

3

Choosing an instrument or method of measurement should be a product of deciding whether the measurement is for summative or formative purposes and whether high or low stakes[2]. There are several measuring instruments available against which mathematics teaching can be measured. MKT, or Mathematical Knowledge for Teaching, is a survey taken by the teacher to measure specialised pedagogical content knowledge[7]. This has been shown to correlate with student outcomes[3]. MQI, or Mathematical Quality of Instruction, is a mathematics-specific lesson observation instrument capturing, among other things, mathematical integrity of explanation, teacher error and precision of language[8]. MQI appears to be more rigorous and detailed in measuring mathematical teaching than value-added models[3]. The TRU (Teaching for Robust Understanding) framework, using the five dimensions of powerful classrooms, aims to support improvement in mathematics teaching as well as help to diagnose areas for professional development[10]. The RATE model of classroom observation (Rapid Assessment of Teacher Effectiveness) has been recently developed as a way to reliably predict student outcomes in mathematics while minimising the need for observer training[12]. Athough not currently widely used, student evaluation questionnaires, such as the one used by Peterson et al[11] can be both reliable and valid sources of teacher evaluation.

IMPLICATIONS: Instruments are available to measure the quality of maths teaching, choice of these should depend on the purpose of measurement

4

Teacher effectiveness as measured by pupil outcomes is not correlated with teacher qualifications or experience[5], even though these measures are very often used to make pay or retention decisions[1]; professional progression or performance decisions for teachers are often based on inadequately reliable measures of effectiveness[2]. Research suggests higher quality teachers and teaching are more likely to occur in the most affluent schools[3] and there is a strong relationship between teacher resources and the quality of their instruction as measured by the MQI[9]. Value-added measures are also related to the resources that students bring to the classroom, for example language skills and support from home[3].

IMPLICATIONS: Current decisions about pay or progression are often made using inadequate measures, such as experience or level of education

School affluence, teacher resources and pupil resources all tend to have a positive effect on student attainment

Lucy Rycroft-Smith, 2017

REFERENCES

1. Cantrell, S., & Kane, T. J. (2013). Ensuring fair and reliable measures of effective teaching: culminating findings from the MET project's three-year study. Seattle, WA: Bill & Melinda Gates Foundation, Measures of Effective Teaching Project.

2. Coe, R., Cesare, A., Higgins, S., & Major, L. E. (2014). What makes great teaching? Review of the underpinning research. The Sutton Trust.

3. Hill, H. C., Kapitula, L., & Umland, K. (2011). A Validity Argument Approach to Evaluating Teacher Value-Added Scores. *American Educational Research Journal*, 48(3), 794–831.

4. McCaffrey, D. F., Koretz, D., Lockwood, J. R., & Hamilton, L. S. (2003). Evaluating value-added models for teacher accountability. Santa Monica, CA: RAND.

5. Schacter, J., & Thum, Y. M. (2004). Paying for high and low-quality teaching. Economics of Education Review, 23, 411–430.

6. Strong, M., Gargani, J., & Hacifazlioglu, O. (2011). Do We Know a Successful Teacher When We See One? Experiments in the Identification of Effective Teachers. Journal of Teacher Education, 62(4), 367–382.

7. Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? Journal of Teacher Education, 59(5), 389–407

8. Hill, H. C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G. C., Sleep, L., & Ball,D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. Cognition and Instruction, 26(4), 430–511.

9. Cohen, D. K., Raudenbush, S., & Ball, D. L. (2003). Resources, instruction and research. Educational Evaluation and Policy Analysis, 25, 119–142.

10. Schoenfeld, A. H., & the Teaching for Robust Understanding Project. (2016). An Introduction to the Teaching for Robust Understanding (TRU) Framework. Berkeley, CA: Graduate School of Education.

11. Peterson, K.D., Wahlquist, C. & Bone, K. Student Surveys for School Teacher EvaluationJournal of Personnel Evaluation in Education (2000) 14: 135.

12. Gargani, J., Strong, M. (2014)Can We Identify a Successful Teacher Better, Faster, and Cheaper? Evidence for Innovating Teacher Observation Systems, Journal of Teacher Education

CAMBRIDGE UNIVERSITY PRESS

UNIVERSITY OF CAMBRIDGE Faculty of Mathematics

UNIVERSITY OF CAMBRIDGE Faculty of Education

Cambridge Assessment