Submission: Response to the Review of the National School Reform Agreement – Interim report

Dr Greg Ashman

I broadly welcome the interim report of the Productivity Commission. There is much that could be written about education in Australia and I am pleased that the Commission recognises initial teacher education as an area of concern.

I also wish to highlight that I am strongly in favour of producing a bank of lesson plans that teachers can draw from rather than have to create everything from scratch, addressing one of the contributing factors to teacher workload, a key concern of the interim report. However, the Grattan Institute is pursuing this agenda and so I will simply state that I endorse it. Instead, in this submission, I will restrict my comments to two main areas.

The first is a suggestion for collecting data that would shed more light on what is happening is real Australian classrooms and the second casts doubt on one of the approaches that is described favourably in the interim report.

The disruptive elephant in the room

The Organisation for Economic Co-operation and Development (OECD) coordinates the Programme for International Student Assessment (PISA), a series of assessments in reading, mathematics and science that fifteen-year-olds in a number of states participate in on a three-year cycle. It is well-known that Australian performance on these assessments is in long-term decline.

Perhaps less well known is that as part of this programme, PISA collects survey data on students' experiences in the classroom. Much of this data is seemingly at odds with conventional wisdom. For example, PISA data shows a negative correlation between the use of 'enquiry learning' in science lessons and performance on the science component of PISA assessments (OECD, 2016), despite enquiry learning being very much in vogue among education academics and bureaucrats. This is not a surprising finding for anyone who has been following the relevant cognitive science (see e.g. Kirschner, Sweller, & Clark, 2006) but it does illustrate that received opinion in education is rarely informed by cognitive science, a point we will return to later.

The survey of students also captures details of the classroom climate that students are exposed to. Students are asked a series of questions about how often, 'There is noise and disorder,' in the classroom or, 'The teacher has to wait a long time for students to quiet down,' and so on. From these questions, PISA constructs an, 'index of disciplinary climate'.

In 2018, Australia ranked 69 out of the 76 countries surveyed (OECD, 2019) in this index of disciplinary climate. New Zealand, ranked 66, was the only other major anglophone country at the same end of the ladder as Australia. The United States ranked 28 and the United Kingdom, 24. This was not a surprising result. In the previous 2015 round of PISA, Australia ranked 63 out of 68 countries on the same measure (OECD, 2016).

Clearly, something is going very wrong with behaviour in Australian classrooms. However, my experience as a teacher suggests this is a problem that teachers, bureaucrats and

academics conspire to ignore. Teachers fear being labelled a bad teacher if they report behaviour problems. It is assumed that they are not planning interesting enough lessons or have not formed good enough relationships with students.

Bureaucrats and academics have imbibed a romantic ideology that stretches at least as far back as the French Philosopher, Jean-Jacques Rousseau and his 1762 book, *Emile, or On Education*. This philosophy views children as innocent and blameless, rather than human and flawed. All misbehaviour is the result of contact with the corrupt adult world and the failure of adults to meet a need that a child is struggling to articulate. This philosophy permeates all levels of education and it is quite common to hear a struggling teacher, at their wit's end trying to manage a challenging class, be uselessly advised that, 'All behaviour is communication'.

There is much that can be done to improve behaviour in the classroom (see e.g. Ashman, 2016). Many of these measures are simple and cheap, such as sitting students in a seating plan designed by the teacher to minimise the potential for distractions and creating common classroom routines. Such strategies act on the antecedents of poor behaviour. Once poor behaviour develops, there are other strategies teachers can use in a graduated and proportionate scale, up to and including the eventual removal of a student from the classroom. Unfortunately, the ideological landscape of education questions whether managing students in this way is legitimate and a human-rights based discourse has developed that opposes any form of 'exclusion'.

It is in this context that we should read statements in the interim report such as:

Student wellbeing affects the wellbeing and effectiveness of educators. In 2018, 24 per cent of Australian lower secondary teachers reported experiencing a lot of stress in their work — 25 per cent reported modifying lessons for students with special needs as a source of stress and 28 percent reported maintaining classroom discipline as a source of stress (Thomson and Hillman 2020, p. 28). Hunter and Sonnemann (2022, p. 16) found that 74 per cent of teachers in 2021 reported that not enough support for struggling students with complex needs was a major issue affecting their time to prepare for effective teaching. For school leaders, the mental health issues of students was in the top four main sources of stress in 2019 and 2020 (See et al. 2022, p. 38)

It is not a minor point to highlight that the wellbeing of students — a key priority of the interim report — is affected by the poor behaviour of other students and this likely forms a vicious circle with students who are affected by others also developing maladaptive behaviours, compounding the problem. This is not just an issue of teacher retention and recruitment, it is about the quality of school life for students.

Recommendation 1 — a national classroom climate survey

I propose we replicate the PISA survey questions as a series of questions students answer after completing one of their NAPLAN assessments. This would be a simple and cheap way to collect data highly relevant to students' wellbeing and classroom experiences. It would enable us to identify pockets of good practice, such as when a school facing challenging circumstances has an especially positive classroom disciplinary climate. Other schools could then learn from this and potentially adapt their practice.

Quality teaching rounds is not the answer

The next point is perhaps less pressing but it is indicative of the sandy foundations that support the edifice of education and education research. The interim report states:

"While they do not offer the same degree of recognition, Quality Teaching Rounds — where teachers work together in small groups to analyse and improve their practice — have been found to have significant positive effects on teaching quality and student academic achievement."

I am familiar with the inexplicably popular Quality Teaching Rounds approach and in my view, the writers of the interim report are unduly optimistic.

If we trace its lineage back through the Queensland School Reform Longitudinal Study (Lingard et al., 2001), we find the origins of Quality Teaching Rounds in the work of U.S. researcher Fred Newmann and his colleagues (Newmann, Marks, & Gamoran, 1996). Newmann's represents a 'constructivist' approach to teaching — constructivism as a theory of *learning* is widely accepted but the idea that it has implications for teaching methods is controversial (see Mayer, 2004). As such, it is the kind of approach to teaching that PISA has been busy finding evidence against and that is at odds with basic findings from cognitive science (Kirschner, Sweller, & Clark, 2006). It therefore does not appear promising.

So, where have the positive results come from?

Quality Teaching Rounds researchers have constructed a rubric for supposedly measuring the quality of teaching based on constructivist principles. This rubric captures key constructivist notions such as that students need to be self-directed and demonstrate supposedly higher order thinking. In one study, teachers trained using this rubric were then observed against it. Following training, they were more aligned with the rubric. Given that the rubric supposedly measured the quality of teaching, the researchers — hey presto! — claimed that the training improved the quality of teaching (Gore et al., 2017). As is apparent, this is entirely circular logic. What if the rubric did not capture the essential characteristics of more effective teaching?

In a more recent study, the researchers finally decided to measure the impact on student outcomes. This led to widely trumpeted headlines that Quality Teaching Rounds improved mathematics performance (Gore et al., 2021). What is rarely mentioned is that mathematics was not the only outcome measure. Data was also collected on outcomes in reading and science. These showed no effect for Quality Teaching Rounds. Data was also collected for a range of noncognitive outcomes. For one of these — a 'Quality of School Life' survey — a statistically significant result was found in favour of one of the control conditions that Quality Teaching Rounds was being compared against.

Moreover, there were *two* different versions of Quality Teacher Rounds. The first was led by the researchers themselves and the other was led by people the researchers had trained. The positive effect on mathematics was *only* found for the researcher-led version of Quality Teaching Rounds. So, even if we believe the programme really did have a positive effect on mathematics performance, there are going to be significant problems with scaling this up to large numbers of schools. The researchers cannot personally lead every implementation.

But it may not be a real effect anyway. There is a 1 in 20 chance of obtaining a statistically significant effect and so if we make enough different measurements, we will eventually find one — a phenomenon known as 'p-hacking' to researchers. Add to this the fact that educational trials cannot be 'blinded' in the same way as, say, pharmaceutical trials. This means that teachers and students tend to know they are in the treatment group and this can produce expectation effects similar to the placebo effect. This is why most education trials of this kind tend to find positive effects for the intervention (see discussion of 'program-based studies' in Zhang et al, 2022).

Recommendation 2 — ditch the references to Quality Teaching Rounds

It will make the final report more credible.

References

Ashman, G. (2016) Classroom Management. *The Truth About Teaching.* Sage.

Gore, J., Lloyd, A., Smith, M., Bowe, J., Ellis, H., & Lubans, D. (2017). Effects of professional development on the quality of teaching: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and teacher education*, 68, 99-113.

Gore, J. M., Miller, A., Fray, L., Harris, J., & Prieto, E. (2021). Improving student achievement through professional development: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and Teacher Education*, 101, 103297.

Kirschner, P., Sweller, J., & Clark, R., E. (2006) Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41:2, 75-86, DOI: 10.1207/s15326985ep4102_1

Lingard, R. L., Ladwig, J., Mills, M. D., Hayes, D., Luke, A., Gore, J., & Christie, P. H. (2001). *The Queensland school reform longitudinal study: A strategy for shared curriculum leadership. Teachers' manual*.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning?. *American psychologist*, 59(1), 14.

Newmann, F. M., Marks, H. M., & Gamoran, A. (1996). Authentic pedagogy and student performance. American journal of education, 104(4), 280-312.

OECD (2016). How schools and teaching practices shape students' performance in and dispositions towards science. *PISA 2015 Results (Volume II)*. Retrieved from bit.ly/3Tmcnp2

OECD (2019). Disciplinary Climate. *PISA 2018 Results (Volume III)*. Retrieved from bit.ly/3gdMf1B

Zhang, L., Kirschner, P. A., Cobern, W. W., & Sweller, J. (2022). There is an evidence crisis in science educational policy. *Educational Psychology Review*, 34(2), 1157-1176.