Teaching and learning for all? The quality imperative revisited

Robin J. Alexander
University of Cambridge, UK

ARTICLE INFO

Keywords:
Education for all
Monitoring
Quality
Pedagogy
Indicators
Measures

ABSTRACT

The 2014 EFA Global Monitoring Report (GMR) is the second with 'quality' in its title but the only one in the series whose title explicitly highlights teaching and learning. While GMR 2014 assesses progress towards the six EFA goals with particular reference to the quality of teachers and teaching, this paper considers progress within the methodology of the monitoring process itself. EFA indicators can attend to only a limited range of variables, and proxies are inevitable. Yet with the post-2015 EFA agenda in view it is essential to ask whether what is truly transformative in teaching and learning has been adequately captured in the GMR monitoring process, the literature on which it has drawn, and the recommendations it has produced. The paper argues for a more radical and creative approach to the defining and use of indicators and arguments that despite pedagogy's pivotal role in generating educational quality, it remains the missing GMR ingredient. The problem is both conceptual and evidential, and the paper argues for a more inclusive, less top–down use of the available research in order to bring into EFA and GMR discourse evidence on teaching and learning that can really make a difference.

© 2015 Published by Elsevier Ltd.

Quality: now you see it …

Like its predecessors, the 2013 – 14 Global Monitoring Report Teaching and Learning: achieving quality for all – hereafter GMR 2014 is impressive in the scale of its evidence, the progress it documents, the warnings it issues, and the humanity of its endeavour.

Quality has been an EFA goal since the 2000 Dakar framework declared it to be 'at the heart of education' and a fundamental determinant of student enrolment, retention and achievement; while, along with quality, learning featured a decade earlier in no fewer than three of the six Jomtien goals. Yet despite these early emphases, quality in the global monitoring reports, and quality in teaching and learning in particular, have since then been surprisingly elusive. In part this may have reflected a preoccupation with those EFA goals whose urgency has seemed the more pressing because their pathology and progress are readily computed. With 57 million children still out of primary school, half of them in 32 countries suffering conflict, and only 13 out of 90 countries likely to achieve universal primary school completion by 2015, we understand why this is so. Numbers offer headlines and dramatic immediacy, ‘Quality’ does not.

Paradoxically, quality may also be elusive because it is ubiquitous. For instance, a consistent argument in the GMRs has been the inseparability of quality from equity, because until an education system is equitable in terms of access, enrolment, gender parity, retention and completion it can hardly be described as being of good quality, even if for some children, in some schools, the experience of learning is rewarding and high standards are achieved. We are justifiably disturbed by the finding of GMR 2014 that while the richest boys may on present trends achieve universal primary education by 2021, the poorest girls will not catch up until 2086. Quality for some is not education for all.

Indeed, quality pervades all six EFA goals. The first GMR called quality a 'composite goal' and one of the strengths of these annual reports is that though each of them has had a specific theme – gender, literacy, early childhood, governance, the marginalised,
conflict, quality, inequality – each has begun by tracking progress towards all six goals as a reminder of the way they are intertwined and must be simultaneously pursued if EFA is to be achieved.5

But quality’s very pervasiveness may have encouraged the view that it requires no further elucidation. So it becomes all the more important to examine how quality has been handled in the EFA monitoring process and how this ‘composite’ goal has been translated into working indicators and measures in the two GMRs – 2005 and 2014 – which have included quality in their titles and remits, for these, post hoc if not a priori, may reveal the definition we seek. Having uncovered that definition, and mindful of the pedagogical orientation of GMR 2014, we can then apply three tests:

- Does the account of quality in EFA attend to what in teaching and learning really matters?
- Are the classroom processes and outcomes that are truly transformative for our children adequately captured in the EFA goals, objectives and targets, the EFA monitoring indicators and measures, and the evidence on which EFA thinking and policy draw?
- If not, what are the implications for the UN’s education mission after 2015, and if learning is to be a target, how should it be defined, indicated and assessed?

In addressing these questions I first return to the analysis I was invited to undertake for the UK Department for International Development (DFID) in 20076 during a period when I was making annual visits to India in connection with the Government of India’s ambitious EFA initiative, Sarva Shiksha Abhiyan (SSA), and its predecessor the District Primary Education Programme (DPEP). The choice of title for this paper should now be clear. In the sense that it re-engages with quality, teaching and learning, GMR 2014 revisits GMR 2005, The Quality Imperative7; and this first revisiting allows a second: a re-assessment of my earlier concerns about how quality, teaching and learning have been handled in the GMR process as a whole.

One of those concerns was the striking neglect of pedagogy despite the fact that pedagogy is at the very heart of education and without pedagogy discussion of educational quality makes little sense. Another was the gulf between the evidence on both quality and pedagogy cited in the EFA GMRs and the much larger body of evidence about these matters that appears in the research literature: one world but two discourses. To counter these tendencies I shall end my paper with an example showing how the EFA movement and its post-2015 successor could increase their effective purchase on the declared priority of advancing quality in teaching and learning if they were prepared to foster a more inclusive discourse and consult a less exclusive literature.

Input, output, proxies and process

Here, briefly summarised, are the problems I identified when I examined EFA and contingent literatures on quality published up to 2007.

First, the quest for indicators and measures of quality produced an understandable preoccupation with input and output – pupil/teacher ratio, balance of male and female teachers, balance of trained and untrained teachers, expenditure per pupil as percent of GDP, net enrolment ratio, adult literacy rate, survival rate to grade 5 – but this was at the expense of indicators of process. Output is in part determined by process but is not synonymous with it. Second, when attempts were made to plug the gap, the identified process elements appeared to reflect not teaching and learning as either experienced or researched but those few random aspects of classroom life that were deemed measurable, regardless of whether they had the significance that their selection implies. Hence, for example, the foregrounding of learning time, time on task and class size.

Third, the very act of isolating such aspects validated them in the eyes of those – governments, administrators, donors – who had the money and power to make them matter, and set in train policies for embedding them ever more exclusively, whether or not this response was justified by the evidence. In this way, the monitoring distorted both what it monitored and the decisions and interventions to which it led. By way of illustration of the risky consequences of this approach we might note that in Lockheed’s and Verspoor’s influential 1991 World Bank cost-benefit analysis of investments for improving primary education in developing countries, pre-service teacher education and midday meals were rejected as ‘blind alleys’.8 Today we take a very different view of the efficacy of both interventions.

Fourth, in an attempt to engage more comprehensively with process, some frameworks posited unashamedly qualitative variables such as ‘high expectations’, ‘strong leadership’, ‘positive teacher attitudes’, ‘appropriate use of language’, ‘committed and motivated teachers’, ‘appropriate teaching and learning materials’, ‘meaningful assessment’, ‘effective management of physical assets’ and the ubiquitous ‘active teaching methods’ and ‘child-friendly environment’.9 But each of these modifiers – high, strong, positive, appropriate, committed, meaningful, effective, active, child-friendly – lacks objective meaning and is open to many interpretations, not just across cultures but also within them, while the overall selection is no less arbitrary notwithstanding its abundance of adjectives.

Fifth, in the absence of watertight measures, compensatory use was made of proxies. ‘Survival rate to grade 5’, as the proxy indicator of quality in the EFA Education Development Index (EDI), is a prominent example.10 This approach is not confined to EFA. Many governments, and certainly the world’s media, treat the performance of a sample of 15 year olds in the PISA tests at a single moment in their educational journey as a valid and reliable measure of the performance of entire education systems. Some proxy.

As a not entirely flippant aside I find the use of ‘survival’ in this context bizarre as well as evidentially ambiguous. ‘Survival’ allows two very different takes, one of them suggesting that education is to be endured rather than enjoyed: (i) ‘How good was your education?’ ‘Excellent: I survived to grade 5.’ (ii) ‘How good was your education? ’ ‘Terrible: I survived to grade 5 but then could take no more and left school.’

This brings me to three overarching problems, which like the tendencies summarised above apply no less in 2014 than in 2007, when I first itemised them.

Quality: a mantra in need of definition

First, there was – and still is – a conspicuous lack of precision in the use of the keyword ‘quality’ itself. Though ‘quality’ is often used quasi-adiectively, as in ‘quality healthcare’, ‘quality teaching’, ‘quality learning’ and so on, it is actually a noun. The adjectival use of ‘quality’, as in ‘quality education’, is no more than a slogan, offering limited purchase on what quality actually entails. But even when used as a noun, ‘quality’ is multi-faceted, for it can mean an attribute – as in ‘the qualities we look for in a teacher’ – or a degree

---

5 The previous GMRs, from 2002 to 2012, are listed on p (iv) of GMR 2014.
6 Alexander (2008a).
8 Lockheed et al. (1991), 87.
of excellence, as when we say teaching is of outstanding quality, in which case ‘outstanding’ needs to be defined. So ‘quality’ – as in Teaching and Learning: quality for all – can describe, prescribe or evaluate.

In the debate about quality in EFA this basic distinction has too often been blurred. That is to say, some have been happy to use supposed indicators of quality in teaching and learning – quality in the sense of a standard to aim for – without adequately exploring and describing those qualities or attributes of which teaching and learning are actually constituted. When we favour prescription over description we risk producing a prospectus for quality which is arbitrary or biased. So I suggest that the task of improving the quality of teaching and learning requires closer attention to the description and analysis of quality and rather less to soundbites like ‘child-friendly teaching’ and ‘active learning.’

Indicators and measures: not the same

The second overarching problem is a confusion between indicators and measures. The terms are frequently treated as interchangeable when they are not. Here I concede that I may well be in a minority within the EFA community (though not, I am happy to say, outside it) for at a 2014 London seminar of the great and good in development education an indicator was defined ex cathedra as ‘a precise metric from identified databases that assesses if a target is being reached’, and nobody thought or dared to disagree. But this definition is stipulative and context-bound and I prefer to argue that if we have the luxury of two terms we should not squander the clarity and nuancing this allows by treating them as synonymous.

Thus, measures measure, indicators indicate: they do different jobs. A measure is a procedure, device or unit for measuring and is irrevocably tied to quantity. An indicator is a more complex and variable clue about whether something is happening and if so to what extent. Approaching clouds indicate the imminence of rain but they do not guarantee it and they certainly do not measure rainfall. A noisy classroom may indicate lack of student concentration but it does not conclusively prove it, still less measure the precise balance of student attention and inattention; indeed, there are those who say that a noisy classroom indicates active learning, but that’s another story.

Take, as one familiar and ostensibly absolute measure of learning in school effectiveness studies, time on task – or what, to distinguish it from available instructional time, Hattie calls engaged time.11 This is often no more than an indicator, and a less-than-conclusive one at that, for it depends only up to a point on objectively-measured time and rather more on an inferential response to students’ observable behaviour. A student who appears to be attending to the teacher, or reading, and hence ‘on task’, may in reality be day-dreaming; or if not wholly off-task then (as happens to all of us) concentrating for only some of the time apparently committed, and with only some of his or her available attention. Who then, from the coding categories offered by conventional classroom observation schedules, is able to calibrate the proportion of actual instructional time when in Hattie’s terms the student is not only engaged but also learning?

These difficulties sustain Gage’s claim that time on task is a ‘psychologically empty quantitative concept’.12 As it happens, Hattie’s survey of 800 student achievement meta-analyses found time on task also to be a poor indicator of both learning and attainment,13 for what matters is the nature of the engagement, not its quantity. Unpacking this relationship further, my own observation of teaching across cultures required me to differentiate the related measure of pace, which in UK school inspections is treated as monolithic and unproblematic, by reference to five elements:

- **organisational pace** (the speed at which lesson preparations, transitions and conclusions are handled);
- **task pace** (the speed at which learning tasks and their contingent activities are undertaken);
- **interactive pace** (the pace of teacher-student and student-student exchanges, and contingent matters such as maintaining focus and the handling of cues and turns);
- **cognitive or semantic pace** (the speed at which conceptual ground is covered in classroom interaction, or the ratio of new material to old and of task demand to task outcome);
- **learning pace** (how fast students actually learn).14

Time, as the above study also pointed out, ‘is a value, not merely a unit of measurement’.15 Of course, the methodological challenge of mental/behavioural inference is at least as old as the discipline of psychology, and time on task is just one of its manifestations. Small wonder that some have chosen to measure learning by focusing exclusively on behaviour qua behaviour. That the problem is at once methodological and conceptual is amply demonstrated in Lefstein and Snell’s more recent take on the problem of pace as identified above.16

Notwithstanding all this, but for reasons we know and understand, what can be measured is privileged in policy circles over what cannot. But this is a dubious and indeed pyrrhic elevation if what is measured has limited indicative power and what is important is marginalised or ignored. Some indicators can be translated into measures, some cannot, but let’s talk about the full spectrum of what needs to be indicated before we start talking about measures. I accept that national education systems, and international education efforts such as EFA, entail massive expenditure and huge populations and therefore require metrics that are as precise as possible and cannot be content with high-inference indicators. But if this imperative excludes what is most important then we have a problem.

To keep open the prospects for engaging with what really matters in teaching and learning I believe that we should sharpen rather than blunt the distinction between indicators and measures, treating the identification of indicators as the first step in the formulation of measures. If it is indeed the case that much that is essential to the quality and outcomes of learning can be indicated but not be measured, we should not arbitrarily exclude such attributes or grasp at proxies which may be conveniently measurable but barely relevant. Instead, we should leave the unmeasurable indicators in place, develop and refine them in their own terms as qualitative devices for making qualitative judgements, and look for appropriate ways of using them to support our tasks of monitoring, development and improvement.

A more radical and creative discussion of EFA indicators is needed than the GMRs have so far provided; one that proceeds from the ‘quality imperative’ of teaching and learning as they irreducibly are, rather from numerical convenience. Hard data are not necessarily useful data.

The task I have outlined applies as much to the assessment of the learning of individual children as to the monitoring of schools and education systems. We can measure children’s mathematical attainment and certain aspects of their basic literacy development. However, GMR 2014 argues, and its argument is welcome, that

---

12 Gage (1978), 75. For a cross-national and cross-cultural discussion of time as a far from objective pedagogical ‘indicator’, see Alexander (2001), 391–426.
13 Hattie (2009), 184–185; Hattie and Yates, 43.
15 Ibid, 425.
16 Lefstein and Snell (2013).

Please cite this article in press as: Alexander, R.J., Teaching and learning for all? The quality imperative revisited. Int. J. Educ. Dev. (2015), http://dx.doi.org/10.1016/j.ijedudev.2014.11.012
while the so-called basics are essential, the fractured nature of our world and the tragedies of poverty and conflict require schools to promote a global citizenship that addresses

issues such as environmental sustainability and peace-building – which require core transferable skills such as critical thinking, communication, co-operation, problem-solving, conflict-resolution, leadership and advocacy – and the promotion of core values such as tolerance, appreciation of diversity and civic responsibility.\(^\text{17}\)

Here we are firmly in the territory of non-measurable indicators. So we must find other ways to describe and assess children’s learning in this vital area. A single testable target or indicator for ‘learning’ across the board – as is proposed for EFA post-2015 – may not suffice, unless it can be proved that, say, numeracy correlates with tolerance, appreciation of diversity and civic responsibility. Actually, this is not as far-fetched as it may seem for as GMR 2014 reminds us, the entire EFA effort is predicated on evidence that education, and especially literacy, reduces poverty, boosts growth, increases employment prospects, enhances health, reduces child mortality, narrows the gender gap and much else.\(^\text{16}\) Even more to the point, a British review of research on citizenship education showed that the skills in question are most effectively developed when they are embedded in the teaching process rather than merely conceived as outcomes.\(^\text{19}\) and this I also take to be the force of the references in GMR 2014 to critical thinking, communication, problem-solving and so on. This, once again, underlines the need for GMRs to engage with classroom process, for that is where citizenship starts.

**Pedagogy: one thing needed**

This takes me to the third overarching problem predicated on my earlier analysis. Brian Simon, the UK’s most distinguished educational historian, famously asked ‘Why no pedagogy in England?’\(^\text{20}\) and we might ask ‘Why no pedagogy in the GMRs?’ If pedagogy is both the act of teaching and the ideas, values, knowledge and evidence that shape and justify it, if it is what the teacher needs to know in order to make valid, effective and defensible classroom decisions,\(^\text{21}\) and if once access and enrolment have been achieved it is what delivers the learning outcomes towards which EFA is directed, then it should have pride of place in a report entitled *Teaching and Learning: quality for all*. But it does not.

In EFA 2002, repeated in subsequent GMRs, there is a table calling itself an input-process-outcome framework for assessing education quality.\(^\text{22}\) At least process is included: all too often it remains securely locked in its black box. But that is as far as it goes, for in this framework ‘process’ comprises just two elements, ‘school climate’ and ‘teaching/learning’. The school climate indicators – high expectations, strong leadership, positive teacher attitudes, safe and gender-sensitive environments, incentives for good results, flexibility/autonomy – are preconditions or contextual factors rather than processes; and the teaching/learning indicators are confined to ‘sufficient learning time’, ‘active teaching methods’, ‘integrated systems for assessment and feedback’, ‘appropirate class size’ and ‘appropriate use of language’.

Apart from the fact that these indicators display, in their use of adjectives like ‘high’, ‘strong’, ‘positive’, ‘sufficient’, ‘active’ and ‘appropriate’, the problem of prescription in the guise of description that I referred to earlier, and uncalibrated prescription at that, most of them are also about context and conditions rather than processes. Only ‘active teaching methods’ and ‘appropriate use of language’ come close, but without further explication these do not amount to much.

In fact, the striking feature of the GMRs is that they do not so much engage with pedagogy as circle around it. Like knowledge itself, pedagogy is a very deep pool. Perhaps UNESCO is afraid of falling in.

**Quality and pedagogy: have the GMRs progressed?**

What has changed during the GMR cycle? I should say immediately that while GMR 2005 was confined to the indicators that have been a constant since Dakar, it was an exception to some of the tendencies I have mentioned. It reviewed definitions of quality from Jomtien, Dakar, the UN Convention on the Rights of the Child and elsewhere, comparing humanist, behaviourist, critical, indigenous and adult education approaches. It also took us back to the 1996 Delors report, *Learning: the treasure within*, whose simple but powerful distinction between learning ‘to know’, ‘to do’, ‘to live together’ and ‘to be’ deserves to be revisited.\(^\text{23}\)

All this was timely and helpful. However, GMR 2005 then proposed a ‘framework for understanding education quality’ in the hope of combining and reconciling the differences which its discussion had exposed.\(^\text{24}\) In fact, apart from juggling the boxes and providing a more detailed elaboration of contextual factors, and in spite of the excellent accompanying discussion of the nature of quality, the quality framework in GMR 2005 was not very different from that in GMR 2002, and its account of teaching and learning – which it revealingly renamed ‘inputs’ rather than ‘process’ – was almost identical. Learning as an input? Only if you view teaching as no more than transmission.

The chronology is interesting, too. GMR 2002 offered ‘a framework for assessing education quality’ while three years later GMR 2005 gave us ‘a framework for understanding education quality’.\(^\text{25}\) Surely it should have been the other way round, for you cannot assess something without first understanding it. Does this back-to-front chronology illustrate a wider tendency in EFA monitoring, I wonder?

What happens when we fast-forward to GMR 2014? Here I can find no exploration, comparable to that provided by GMR 2005, of what educational ‘quality’ means. I assume that this is because it would look odd still to be debating such matters after ten reports and just one year before GMR 2015, the final report in the series. So in the assessment of progress towards EFA Goal 6, quality is characterised by the ‘key indicators of pupil/teacher ratio at the pre-primary, primary and secondary stages, the continuing teacher gender imbalance and the availability of textbooks.’\(^\text{26}\) To these is added a section on the need to strengthen international and regional assessments.\(^\text{27}\)

All of these are important, but are they sufficient? And where, once again, are the processes of teaching and learning which GMR 2014 itself acknowledges are so vital to the EFA effort? ‘Strong

\(^{\text{17}}\) GMR 2014, 295.

\(^{\text{18}}\) GMR 2014, 144ff.

\(^{\text{19}}\) Deakin Crick et al. (2005).

\(^{\text{20}}\) Simon (1981) His critique was revisited a quarter of a century later in Alexander (2004), ‘Still no pedagogy?’.\(^\text{16}\)

\(^{\text{21}}\) Alexander (2000b), 47. The full definition is: ‘Pedagogy is the act of teaching together with its attendant discourse of theories, values and evidence. It is what one needs to know, and the skills one needs to command, in order to make and justify the many different kinds of decision of which teaching is constituted.’ This definition is more continental European than Anglo-American, and reaches back to Jan Komensky (Comenius) rather than to those 1960s US/UK curriculum builders for whom curriculum was paramount and pedagogy subsidiary. The neglect of pedagogy in the GMRs may relate to this unexamined legacy as well as to other problems this paper identifies.

\(^{\text{22}}\) UNESCO (2002), 81.

\(^{\text{23}}\) Delors et al. (1996).

\(^{\text{24}}\) GMR 2005, 36.

\(^{\text{25}}\) GMR 2002, 81; GMR 2005, 36.

\(^{\text{26}}\) GMR 2014, 84–89.

\(^{\text{27}}\) GMR 2014, 89–99.
national policies that make teaching and learning a high priority are essential’, says the report, ‘to ensure that all children in school actually obtain the skills and knowledge they are meant to acquire.’

Just so. However, in the next paragraph expectations that at last we are getting somewhere are dashed when ‘teaching quality’ becomes ‘teacher quality’ (my italics in each case) and this unexplained but significant shift from act to agent is then consolidated in the report’s detailed discussion of teacher numbers, recruitment, qualifications, subject knowledge, training, retention and governance.

The emphasis on teachers is supremely important. Without teachers there is no teaching, and without good teachers the learning potential of many children will remain untapped. The association between teacher quality and learning outcomes is both self-evident and empirically demonstrable. But what are teachers to teach and how? And on what aspects of their teaching should their training concentrate, and why? And can we answer these questions if the nature of teaching has been inadequately conceived?

GMR 2014 does engage with some of these questions. It emphasises training for pupil diversity, gender parity and children with learning difficulties. It argues the need to compensate for teachers’ poor subject knowledge and the importance of tools for classroom diagnosis and assessment, especially in relation to children at risk. And then, in its crucial seventh and final chapter, ‘Curriculum and assessment strategies that improve learning’ it at last enters the classroom.

So 14 years after Dakar are we there at last? Have we finally reached pedagogy?

Yes and no, but mainly no. The discussions of both curriculum and assessment are, within the limits they set themselves, useful. As I have noted, GMR 2014 departs from the exclusive preoccupation with literacy and numeracy and argues the need for a wider curriculum and transferable skills. However, it sticks to the received view, dating back to the 19th century, that literacy and numeracy are and forever should remain the sole ‘basics’ of education, regardless of time, location, culture or national circumstance.

In this matter, the case for literacy remains exceptionally strong as both a tool for individual empowerment and a lever for social and economic progress, and successive GMRs have convincingly documented its impact in these terms. But, heretical though some may find the thought, the case for continuing to give numeracy parity with literacy is neither proved nor even entertained; the habit of history, it seems, is sufficient justification, and because ‘literacy-and-numeracy’ has become in effect a single curriculum component, numeracy gets a free ride. Thus we are offered a curriculum in which only literacy, numeracy and citizenship are deemed ‘basic’. But where, some beneficiaries of citizenship education might ask, are science or IT? And where, given the reference to transferable skills for citizenship, is the no less compelling evidence on transfer of learning through the arts? Questioning fixed curriculum mindsets is surely as necessary a part of the GMR exercise as revisiting habitual assumptions about what constitutes a valid educational indicator, and if the task is thought to be necessary in rich countries, why not elsewhere?

Commendation with reservation also applies to the treatment of assessment. GMR 2014 breaks new ground in EFA circles (though not elsewhere) by discussing formative as well as summative assessment, or what in the UK is called ‘assessment for learning’. But here the discussion is again frustrated by the GMR’s limited apprehension of pedagogy. Effective assessment for learning is more than the tools, boxes and packs that in this context GMR 2014 recommends from examples in Uganda, Liberia, South Africa, Colombia and India, which indeed their evaluations show to be effective in terms of both diagnosis and outcomes. More fundamentally, assessment for learning is the very stuff of which effective teaching is made: the day-to-day, minute-by-minute observations and interactions through which good teachers constantly monitor children’s learning and progress, affording the feedback which will build on their understandings and probe and remedy their misunderstandings.

On this vital matter GMR 2014, like its predecessors, has little to say. Once again we trip over the black box or meet the timorous figures circling the deep pool of pedagogy. Curriculum prescribed but not enacted; summative assessment but not formative; input and outcome but not process.

Why no pedagogy in the Global Monitoring Reports?

If I am right that pedagogy is the missing ingredient in accounts of educational quality in these global monitoring reports, and that where pedagogic process appears its treatment is confined to random indicators, and that these tend to circle the teaching-learning process rather than engage with it, then as a prerequisite for improving matters in the post-2015 agenda we must urgently ask why this should be so. Why no pedagogy?

One answer is that when the availability and competence of teachers is a major challenge, as the GMRs show that it is, then it makes sense to focus on teachers rather than teaching, invest heavily in teacher recruitment, training and retention, and develop textbooks and classroom materials which in 1960s US parlance are ‘teacher proof’ and will enable even the minimally-trained teacher to do a reasonable job. On that basis, it may be thought that there is more to be gained from providing such materials than advocating more sophisticated and interactive models of teaching, especially in the context of large classes and multi-grade teaching. In these situations, textbooks and TLMs provide a predictable and reliable foundation for the teacher’s work, effective even when the teacher is absent. For, as GMR 2014 reminds us, teacher absenteeism remains a major impediment to EFA.

This argument is persuasive, though we must ask whether it is right for all circumstances and all teachers, and to what extent it should inform the EFA agenda after 2015. Making teaching ‘teacher proof’ may safeguard educational minima and compensate for teachers’ poor training or erratic attendance, but it can be disempowering and, for competent and talented teachers, demeaning.

But is there another explanation for the neglect of process? I think there is, and it resides in the literature and evidence on which, since 2002, the GMRs have drawn. A head count of the 680 or so published sources listed at the end of GMR 2014 reveals that in a report promisingly entitled Teaching and Learning the titles of only 40 of the cited publications – a mere 6 per cent – refer, directly or indirectly, to the report’s claimed focus. A somewhat larger proportion deal with teacher supply, training and retention, and a much larger proportion still are
macro-level national or cross-national studies of education policies, programmes, strategies, governance, funding and outcomes.\(^{35}\)

That apparently skewed citation profile encourages us to dig deeper. A decade ago, a review undertaken for the US National Research Council (NRC) identified three main types of international comparative study in education. Type 1 are large-scale policy-directed statistical studies of educational achievement, expenditure and other matters of the kind that emanate from OECD, the World Bank and the UN. Type 2 are desk-based extrapolations from international data aimed at identifying policy options and solutions (Michael Barber’s three McKinsey reports are pertinent examples). Type 3 include the majority of studies in the published corpus of academic comparative, international and development education. These range from broadly descriptive accounts of individual education systems to the ‘thick description’ of close-grained cross-national and cross-cultural comparative studies of school and classroom life and the forces that shape it.\(^{36}\)

Types 1, 2 and 3 add up to a literature of considerable variety and richness. However, the NRC report adds that while the majority of published comparative education studies are Type 3, and while many Type 3 studies have significant policy applications, it is the Type 1 and 2 studies that receive most of the funding, policy attention and publicity. Meanwhile, Type 3 studies have more limited funding and rarely come to the attention of policy makers or Type 2 reviewers. The NRC report judges this to be deeply unfortunate because Type 3 studies engage with education, teaching and learning as they are enacted and experienced in schools and classrooms to an extent that Type 1 and 2 studies do not and by virtue of their methodology cannot. This neglect of Type 3 evidence reinforces the remoteness of policymakers and the policy process from schooling as it is experienced by teachers and children, and increases the risk that high-cost and high-stakes interventions relating to teachers, teaching and learning may be misconceived or misdirected.

Following the NRC analysis, what we may have, then – not universally or inevitably, but too frequently – is a six-fold problem of evidential selectivity in the corridors of power. First, the preferred evidence is top-down. It reflects the world, the preoccupations, the priorities and the experiences of policymakers rather than those of teachers and children. Second, it may privilege a supposedly international but essentially western perspective over an indigenous one. Third, its view of school and classroom life may be generalised, coarse-grained, unnuanced and perhaps simplistic. Fourth, its understandable pursuit of what can be measured removes from the agenda and consciousness of policymakers those vital aspects of education that quantification cannot access. Fifth, it ignores a substantial tranche of evidence of which, in the interests of competent and democratic policy making, policymakers, or at least their advisers, have a duty to be aware. Sixth, it is self-sealing and self-reinforcing. Reading UK government publications I am constantly struck by the extent to which they refer only to other government publications. Such circularity in evidence, argument and policy is always dangerous, a fortiori in the context of global education.

It would be impolite of me to accuse GMR 2014 of these tendencies, but given what I have said about the balance of published sources listed at the end of the report the possibility at least deserves consideration, for GMR 2014, like all the GMRs, leans more towards Type 1 evidence than Type 3. In doing so, is it missing something important? I think it is.

### Engaging with pedagogy: conceptual and empirical possibilities

I want now to show how in the elusive area of pedagogy such evidential selectivity and imbalance can be avoided and how we can then greatly enhance the debate about the quality of teaching and learning in EFA. I will deal with the matter first conceptually, then empirically.

The genealogy of the teaching-learning framework in GMR 2002 which was modified in GMR 2005 and remains influential in EFA is clear: 1960s US process-product research transmuted into 1990s transatlantic school effectiveness research and domesticated by international agencies like the World Bank.\(^{37}\) It atomises rather than synthesises, includes only what can be easily measured, views teaching as simple transmission and so concentrates much more on the teacher than the learner, and treats culture not as an all-pervasive feature to be handled with care, sensitivity and humility but as just another variable to be factored and crunched.

I would not claim that my own alternative is impervious to critique, still less that it is the one to adopt, but it at least provides a contrast. Striving to develop a framework for the analysis of both quantitative and qualitative classroom data from five very different education systems in Europe, North America and Asia, I started with what I believed were two irreducible propositions about the nature of teaching, as it is in any context: (i) Teaching, in any setting, is the act of using method x to enable students to learn y; (ii) Teaching has structure and form; it is situated in, and framed by, space, time and patterns of organisation, and it is undertaken for a purpose. From these I derived a two-part framework or matrix comprising the act of teaching and the ideas that inform it.

Teaching as act was divided into (i) the act itself, subdivided into the planned learning tasks, the activities and interactions through which tasks are mediated and the judgements by which students’ needs, progress and attainment are assessed; (ii) the form by which units of teaching are bounded (usually the lesson); and (iii) the organisational, curricular, epistemic and temporal elements of its frame. Each of these was then further elaborated to support a mixture of quantitative and qualitative data analysis.\(^{38}\)

Teaching as ideas encompassed values, beliefs, theories, evidence, policies and justifications and in respect of these differentiated three levels or domains: (i) classroom (ideas relating to students, learning, teaching and curriculum that enable teaching on a day to day basis); (ii) system/policy (ideas about schooling, curriculum, assessment and other matters that formalise or legitimise teaching); and (iii) cultural/societal (ideas about community, culture and self that locate teaching).\(^{39}\)

Teaching as act identifies the cross-cultural structural invariants of teaching. This part of the framework has been shown to be not only comprehensive but also as culture-fair as any such cultural artefact can be. At this moment, for example, a research team at Ben Gurion University is using it to analyse Israeli pedagogy.\(^{40}\) Teaching as ideas accesses the cultural variables that shape, breathe life and meaning into, and indeed define these invariants and thus demonstrate the extent to which (as I showed earlier) even the most securely quantifiable of them is culturally loaded.

---

\(^{35}\) GMR 2014, 410–443. I cannot of course claim to have read every one of the cited publications, and am basing this assertion on scrutiny of their titles, authorship and publishing details. But it is an analysis I have applied elsewhere on the basis of full bibliographic search to other ‘Type 1’ publications: see, for example, Alexander (2012).

\(^{36}\) National Research Council (2003), ‘Thick description’, the ethnographer’s credo, is from Geertz (1983).

\(^{37}\) For analysis and critique of this paradigm: Alexander (2001), 26–40; Alexander (2008b), 9–42.


\(^{40}\) For an early paper from this project: Lifstein (2013).

---
In my own cross-cultural studies I found that such ideas concerned not just the nature and purposes of learning, knowledge and teaching – transmission, induction, negotiation, facilitation, acceleration and so on – but even more fundamentally what I called ‘primordial values’ about the relationship of the individual to others and to society which translate into culturally-distinctive classroom routines and patterns of organisation. This could take us into discussion of so-called ‘western’ and ‘non-western’ models of teaching for which there is not space in this paper, except to note that to portray the cultural diversity of teaching and learning as conceived and enacted across 196 nations and thousands of cultures and sub-cultures as a simple choice between ‘western’ and ‘non-western’ is crude in the extreme. Note, too, that this dichotomy makes ‘western’ the default and ‘non-western’ the aberration.42 Edward Said would have had something to say about that.43

There is no way that the current GMR paradigm can capture any of this. What the comparison of these frameworks also signals is another important question: are there universals in teaching and learning that apply across cultures and contexts, or is everything culturally unique? In my own work I strenuously argue that culture and history are the keys to understanding and comparing national education systems. But I also believe from what we know about human development and education across cultures that there is a level at which pedagogic universals can be defined. My complementary frameworks for teaching as act and ideas try to capture these.

Frameworks like those I have exemplified expose the conceptual incompleteness of the input-output models in GMR 2002. Above all, the classroom interaction through which both learning and teaching are mediated is almost absent from the GMR frameworks. Let us therefore stay with interaction, mindful of the NRC’s judgement that policy-directed studies lean too exclusively on Type 1 and 2 research and ignore Type 3 – a judgement which, in relation to the interactive heart of teaching and learning is born out by the bibliography of GMR 2014. Where, then, can we go to plug this gap?

There is a considerable literature on classroom interaction in general and educationally productive talk in particular, but I will mention by way of example just two major sources that have the virtue of being comprehensive, methodologically diverse, cross-cultural, cross-national and rigorously empirical. Further, their publication dates coincide neatly with GMR 2014 and our consideration of the post 2015 agenda.

The first is a collection of research papers arising from an international conference on classroom talk which was convened in 2011 at the University of Pittsburgh, USA, under the auspices of AERA.44 This brought together many of the world’s leading researchers in the areas of pedagogy and linguistics to establish whether, after several decades of research, we have proof of concept that high quality classroom talk not only engages children’s attention and participation – as we have known for a long time that it does – but also raises their standards of achievement in tests of literacy, numeracy and science.

The answer to that question was conclusively affirmative. There is now a critical mass of randomised control studies in different countries showing that high quality classroom talk enhances understanding, accelerates learning and raises measured standards. This finding is also confirmed in Hattie’s synthesis of 800 meta-analyses relating to student achievement in respect of interactive strategies such as reciprocal teaching, peer tutoring, student verbalisation and feedback.45 Such strategies, in Hattie’s words, make children’s learning visible to the teacher and hence amenable to appropriate diagnosis, assessment and intervention (Hattie, 173–8). The quest for indicators of visible learning would be a useful exercise for the team of GMR 2015. ‘Visible’ and ‘measurable’ are not, however, synonymous.

The other study counters the claim that because the research I have cited comes from classrooms in high income countries it cannot fairly be expected to apply in the context of EFA. This second study is a review of research on pedagogy, curriculum, teaching practices and teacher education in developing countries which the UK government’s Department for International Development (DID) commissioned from the University of Sussex.46 Having trawled 489 studies from middle and low income countries, the Sussex team examined 54 of these in depth. While acknowledging the methodological limitations of some of the studies the Sussex team nevertheless felt able to conclude that classroom interaction is the pedagogical key. They highlighted as feasible and proven strategies for effective teaching in these contexts inclusive and supportive communication, varied teacher questioning, informative feedback, building on student responses, student questioning, and other elements of what I call dialogic teaching.47

All the studies in these important US and UK collections, which together include classroom research from high, middle and low income countries, are Type 3. Being Type 3 and engaging with teaching as it happens they show not just that high quality classroom talk makes a difference but how it can be improved. But only a handful of the studies in the Sussex report, and none of those from the AERA symposium, found their way into the vast bibliographies of GMR 2005 and 2014, dominated as these were by Types 1 and 2. In the UK, meanwhile, a major research and development project has been initiated which will capitalise on such evidence to develop, test and evaluate the capacity of talk-rich teaching strategies to close the achievement gap between some of Britain’s most disadvantaged children and their more advantaged peers.48

Conclusion: what is to be done?

And so we return to my initial questions:

- Does the account of quality in EFA attend to what really matters in teaching and learning?
- Are the classroom processes and outcomes that are truly transformative for our children adequately captured in the EFA goals, objectives and targets, the EFA monitoring indicators and measures, and the evidence on which EFA thinking and policy draw?
- If not, what are the implications for EFA after 2015, and if learning is to be a target, how should it be defined, indicated and assessed?

I submit that in respect of the monitoring of quality in teaching and learning in EFA we have a problem which is both conceptual and empirical. Neither quality nor pedagogy are adequately conceived, and some of the world’s most important and relevant evidence on teaching, learning and their improvement has been

44 Hattie (2009).
45 Westbrook et al. (2013).
48 These matters are discussed in a paper first presented at the 2002 AERA Annual conference and later revised for Alexander (2008b), Chapter 4.
50 Resnick et al. (2015).
ignored. Classroom interaction is the most prominent and perhaps crucial aspect of pedagogy, among several, to suffer this fate and the example above shows that its omission is both grave and unnecessary.

**How can this unsatisfactory situation be addressed?**

First, education for the period post-2015 needs a radical and properly informed debate about indicators and measures in relation to the black box, or black hole, of teaching and learning. For classrooms are the true front line in the quest for educational quality. The proper sequence, surely, is not to make do with the odd measure that happens to have featured in a number of school effectiveness studies but to start with a rounded account of the educational process and the purposes it serves, then range comprehensively and eclectically across the full spectrum of relevant research and extrapolate what the evidence shows can safely be regarded as key indicators of quality, and only then proceed to the question of how those indicators that have been shown to have pre-eminent influence on the quality and outcomes of learning can be translated into measures. In all cases, both indicators and measures should resonate clearly with goals. It is all too common for education goals grandly to espouse \( x \) and then signal through what is tested that the true priorities are rather different.

Second, where an indicator has empirical provenance but cannot readily be translated into a simple measure, other ways should be found to keep it in the frame. Under no circumstances should an indicator that peer-reviewed research has shown to be optimal to effective teaching and standards of learning be dropped at this stage merely because it cannot be quantified. We have to find other ways of handling it. We need a more creative and less doctrinal approach to the whole question of indicators and measures, exploiting, as I suggest above, the methodological possibilities that the vocabulary encourages. In any case, the objectivity of quantitative measures is often overstated, while there are established procedures for ensuring inter-judge reliability in the use of so-called subjective assessments. Again, I warn against paradigm wars.

Third, to cover the evidence as it needs to be covered, teams working on the defining and monitoring of quality in education post 2015 should become more relevantly multi-disciplinary than, in the EFA context, they appear to have been thus far.

Fourth, let’s accept that although much Type 3 evidence comes from and relates to high-income jurisdictions and systems, the recent DfID review of research on pedagogy, curriculum and teacher education in developing countries shows that there is now a fair Type 3 corpus from middle and low income contexts too. One of the post-2015 tasks, I suggest, should be to expand that corpus and make it as reliable as possible. In any case, there is sufficient evidence from cross-national studies of teaching and learning, and from Hattie-style meta-analyses of classroom research, to show that there are universals to which in any event we should attend; for example, teacher professional content knowledge (not the same as subject knowledge), the character and degree of cognitive challenge afforded by teacher-student interaction, and the quality of the information conveyed in teacher-student and student-teacher feedback.

Fifth, having identified which processes matter most and having nominated them as essential indicators, we will find from Type 3 research that some of them are more amenable to measurement than may be thought. That goes especially for the teacher-student interaction that lies at the heart of teaching.

Sixth, in light of all this, we need therefore to explore targets and indicators for both learning and teaching. Learning needs a process indicator as well as an outcome one, and on the basis of what we know about the crucial conditions for learning, we might try student engagement. Similarly if teaching has to be reduced to just one indicator, on the basis of what we know about the characteristics of effective teaching from both Hattie’s meta-analysis of studies in high income countries and the 2013 DfID literature review, we might try reciprocity in teacher-student interaction. As it happens, both engagement and oral reciprocity are amenable to measurement, so mine is not a completely hopeless cause. Having said that, they are also susceptible to the same problems of behavioural inference that I discussed earlier in relation to time on task, which underlines the need for continuing caution and vigilance.

Finally, here is the double and troubling dilemma. I and others want teaching and learning as they happen and are experienced to gain the prominence in education after 2015 that they deserve and urgently need if we are to make progress, and because the modality highlights targets, indicators and metrics, that means that we need to start by exploring how far what matters in teaching and learning is amenable to this treatment. But on the basis of the reductionism we have witnessed thus far, the prospect of a single global measure of the quality of teaching applied across all cultural and pedagogical contexts is nothing if not deeply alarming. Quite apart from the totalitarian resonance of such an idea, or the pejorative charge – namely, probability – that the measure would be plain wrong, teaching is a quintessentially local activity; and I say this having argued above that there are observable structural variants in teaching and learning that are encountered across cultures and systems.

This also means that while generalised process quality targets may be volunteered on the basis of what the evidence tells us makes a difference, it is only at classroom level that they can be feasibly monitored. The trick will be to give process and the quality of process the prominence they deserve without allowing the resulting indicator(s) to tyrannise and debase what they purport to advance and improve; and to find a way to add this essential local dimension to development and monitoring processes that are no less essentially about global and national development.

There are many barriers to achieving education for all, but evidence should not be one of them. Quality in teaching and learning is a global imperative. It demands a global community of discourse. I hope that UNESCO and its advisers will approach global education post-2015 with a commitment to make much more inclusive use of the abundant evidence on pedagogy that is now available in order to exert maximum impact on quality where it matters: in the classroom. Stop tiptoeing around the pool of pedagogy. Take the plunge.

References


Alexander, R.J., 2008a. Education for All, the Quality Imperative and the Problem of Pedagogy. CREATE Research Monograph 20. CREATE/IOE, Sussex/London.


Please cite this article in press as: Alexander, R.J., Teaching and learning for all? The quality imperative revisited. Int. J. Educ. Dev. (2015), http://dx.doi.org/10.1016/j.ijedudev.2014.11.012


Lefstein, A., 2013. The rules of pedagogical discourse in Israel: has the time come to break them? Paper presented to the Israel Ministry of Education.


President’s Commission on the Arts and Humanities, 2011. Reinvesting in Arts Education: Winning America’s Future through Creative Schools. PCAH, Washington, DC.


