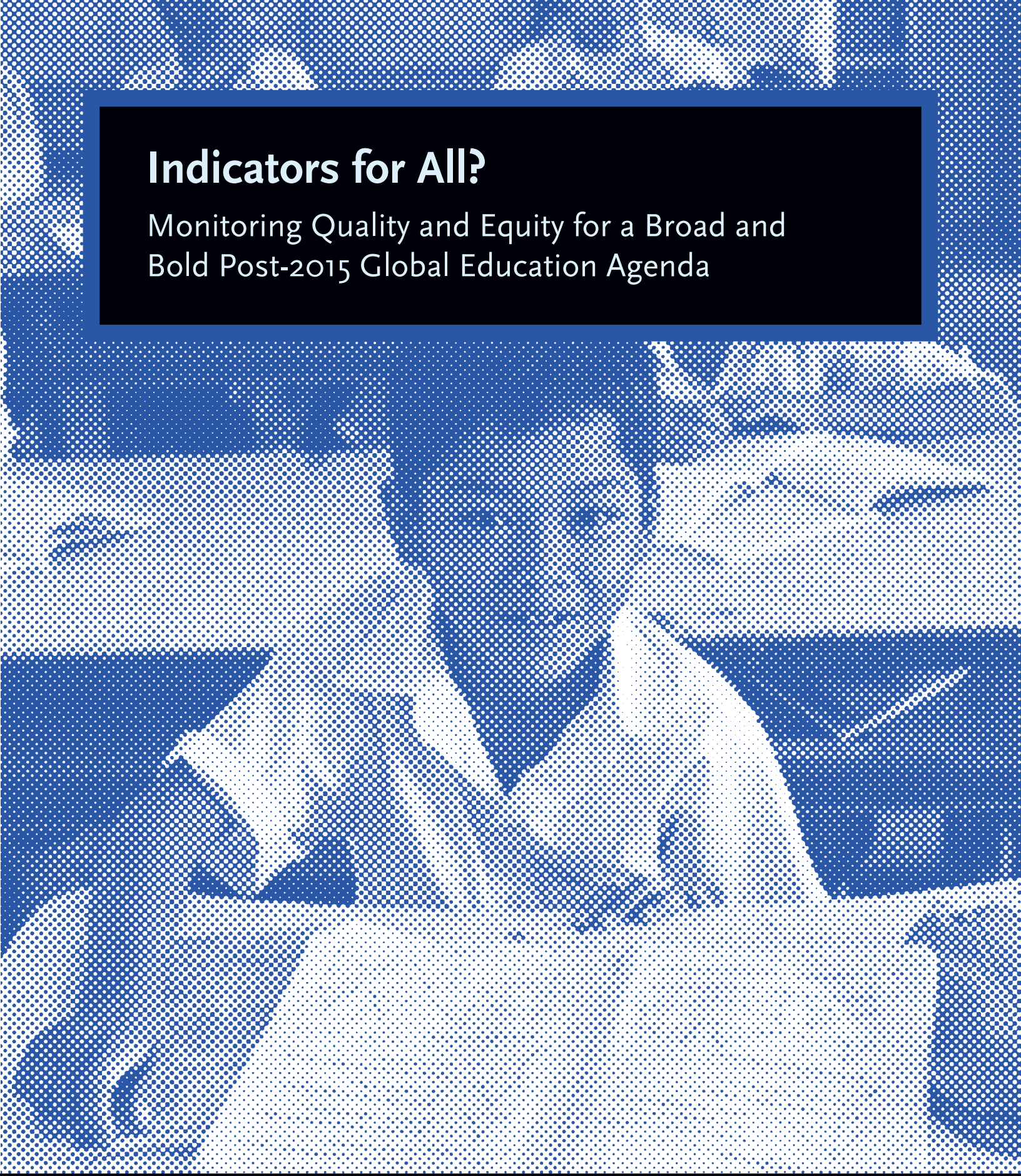
Indicators for All?

Monitoring Quality and Equity for a Broad and Bold Post-2015 Global Education Agenda



OPEN SOCIETY FOUNDATIONS

APRIL 2015

Indicators for All?

Monitoring Quality and Equity for a Broad and Bold Post-2015 Global Education Agenda

Angeline M. Barrett and Tore Bernt Sørensen
University of Bristol

Copyright © 2015 Open Society Foundations

This publication is available as a PDF on the Open Society Foundations website under a Creative Commons license that allows copying and distributing the publication, only in its entirety, as long as it is attributed to the Open Society Foundations and used for noncommercial educational or public policy purposes. Photographs may not be used separately from the publication.

Cover photo: Children in school, Siem Reap, Cambodia. © 2015 Pascal Deloche/Getty.

OPEN SOCIETY FOUNDATIONS

APRIL 2015

CONTENTS

FIGURES

ABBREVIATIONS

ESD	Education for Sustainable Development
EFA	Education for All
IEA	International Association for the Evaluation of Educational Achievement
GCE	Global Campaign for Education
GMR	Global Monitoring Report
OECD	Organization for Economic Co-operation and Development
OWG	Open Working Group
MDG	Millennium Development Goals
NER	Net Enrolment Rate
SABER	Systems Approach for Better Education Results
SDG	Sustainable Development Goals
TAG	Technical Advisory Group
TALIS	Teaching and Learning International Survey
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund

ACKNOWLEDGEMENTS

We are grateful to Terra Sprague, University of Bristol, and Lindsey Horner, University of Sussex, for their constructive and insightful comments on a draft of this paper.

EXECUTIVE SUMMARY

I. INTRODUCTION

This paper sets out to reclaim rights-based thinking on for the current proposals for the education Sustainable Development Goal (SDG). The proposed indicators are intended to contribute to a broad and bold agenda for education quality compatible with the targets set out in two proposals for a post-2015 education goals – the Muscat Agreement (Global Education For All Meeting 2014) and Open Working Group (OWG) proposal (Open Working Group for Sustainable Development Goals 2014). Two targets are addressed, the "relevant learning outcomes" component of the basic education target and the target for qualified teachers.

The paper is organized into two parts. Part I looks back at experience with the education MDG, EFA goals, and understandings of quality within the EFA movement to arrive at a framework for formulating indicators. Suggestions for post-2015 indicators are set out in Part II. The conclusion argues that the development goal should not only monitor what has been achieved against a pre-determined agenda but support stakeholders across all levels to create and implement a broad and bold agenda for education.

PART I: LOOKING BACK: INDICATORS AND QUALITY WITHIN EFA

2. The role of indicators in realizing targets and goals

2a. Development goals and their influence

Development goals can achieve or reflect a consensus; prioritize or boost certain areas of action; act as a planning tool for policy; or act as "a soft form of accountability" (Langford 2012: 12). This last function has intensified over the lifetime of the MDGs in part due to the use of evolving technologies of quantification, such as international learning surveys. The EFA goals and the education MDG have had such influence internationally that they have come to constitute global education policy, which carries the international influence of agendas generated in Western metropolitan centres. Despite the remarkably wide-ranging consultation, this trend is evident in the post-2015 debate and escalates the challenge of setting an agenda that is "supportive rather than confining" for countries (Ahmed 2014: 64).

2b. What does past experience tell us about the impact of indicators?

The education MDG has increased access to a complete cycle of primary education. However, many learners, particularly from disadvantaged groups, benefit very little from schooling. The prioritization of expansion is an instance of an indicator reshaping its parent norm. A clear lesson is that indicators that are only partially fit-for-purpose in terms of how well they capture the meaning of the target can come to displace the

parent target. Statistically robust indicators of what is readily measurable are very often only partially fit-for-purpose. Some measurement experts have suggested using them in combination with less robust indicators but more fit-for-purpose targets to construct a set of indicators that together are fit-for-purpose. Being fit-for-purpose in practice, however, will also depend on monitoring and reporting mechanisms that amplify qualitative indicators, which tend to attract less attention. Indicators are more likely to have traction if they are comprehensible to and valued by educational professionals and civil society advocates of EFA, and can harness the support of wider society. This is called "communicability" or "salience" (Langford 2012: 20). A sustainable development agenda will also require national and sometimes sub-national level indicators and use of indicators for which data is not currently available.

2c. What version of quality should indicators promote?

Fitness-for-purpose of indicators needs to be evaluated not only with reference to the agreed wording of the target but also the human rights that the target aims to realize. They should clarify and elaborate on targets in ways that are consistent with the overarching sustainable development agenda.

2d. Criteria for indicators that create and sustain a broad agenda for quality

Lesson learned from experience with the education MDG and EFA can be distilled into five criteria for indicators that set an implementable agenda for quality and equity in education.

1. Compatible with the right to education
2. Contribute to sustainable development
3. Relevant to diverse contexts
4. Equity
5. Fit-for-purpose
6. Communicable and salient

3. Conceptualizing and targeting quality in EFA

Section 3 looks back at the history of how quality has been conceptualized in the EFA movement from 1990 onwards as a source of ideas for targeting quality post-2015.

3a. Rights-based understandings of quality

Since 1990, theorists and rights-based organizations have developed a broad and bold agenda for quality derived from Convention on the Rights of the Child and other treaties. Their frameworks encompass benefits to learners and the societies in which they live. For individuals, instrumental benefits include meeting basic learning needs particularly skills in literacy and numeracy; and a broad range of knowledge, skills and values in the

cognitive, social, emotional and creative domains. It develops capabilities for contributing towards national development goals and positive participation in society, including leadership and citizenship knowledge and skills related to gender, health, nutrition, peace and respect for the culture of others. These outcomes are realized through classroom and school processes that are directly experienced by the learner, processes that recognize and respect what the learner brings – her socio-cultural background, identity and prior knowledge; that engage with the learner's community; that ensure the learner is well-nourished and ready to learn; that create a safe and healthy learning environment for girls and boys. Equitable school and classroom processes are enabled through a series of inputs: adequate physical infrastructure, well-trained qualified teachers, relevant curriculum and learning materials, participatory governance and management, and accurate assessment of learning.

This body of work presents a broad and bold vision for education quality. The Muscat and OWG proposals both expand this agenda further through integrating it with the sustainable development agenda. But they also crop the agenda by containing it within targets for enrolment and measurable learning outcomes; less measurable outcomes of knowledge, skills, values and attitudes for sustainable development and peace/citizenship; and inputs of infrastructure and teachers. Solidifying a broad and bold agenda through constructing indicators for a broad vision of quality is challenging, not least because the benefits of a quality education are situated within diverse geographies, economies and cultures.

3b. Defining system level indicators

Pigozzi's (2008) rights-based framework identified two levels of organization that enable learning, the level of the learner and the level of the learning system. Some proposals generated by the post-2015 debate include system level conduct indicators, such as "nine years of free and compulsory basic education in legal/institutional frameworks" (Post-2015 Education Indicators TAG 2014) but system level indicators have not been developed systematically across the targets.

System level indicators are not without difficulty. OECD programmes such as PISA, the World Bank's SABER programme and high profile research by large consultancy companies such McKinsey & Co. have attempted to benchmark education policies. They have been heavily critiqued for disregarding local, national and regional diversity in institutional arrangements and the fact that education systems are embedded in socio-cultural environments with distinctive traditions, norms and practices (Alexander 2000; Goldstein 2004; Nardi 2008).

In suggesting the use of process indicators, this paper seeks to formulate them in ways that support context-sensitive problem solving and capacity building across different levels by allowing for adaptation and a variety of approaches within an overarching rights-based sustainable development agenda.

3c. A framework for designing indicators

We take the rights-based vision for education quality together with the distinctions between levels and types of indicators presented to start constructing a framework for designing indicators (see figure 3). At this stage, we do not formulate indicators but simply match up some of the characteristics identified with a quality education with the different types of indicator (outcomes, inputs and processes) and suggest the appropriate level (international or national) for determining indicators.

PART II: LOOKING FORWARD: RIGHTS-BASED INDICATORS FOR THE EDUCATION SDG

The second part of this paper focuses on indicators for two of the post-2015 targets.

4. Indicators for "relevant learning outcomes"

The Muscat and OWG targets are concerned with learning outcomes of various kinds. In this paper we focus on relevant learning outcomes for the basic education target but interpret relevance with reference to the other targets that speak to educational outcomes. We view relevance as having socio-cultural and socio-economic dimensions concerning preparation for life in local and national contexts and participating in the benefits of globalization.

4a. Learning outcomes as indicators of equity

Relevance has not been addressed by much of the existing work on post-2015 indicators. Measures of learning are commonly treated as unproblematic indicators of quality (e.g. Filmer *et al.* 2006) and important indicators of equity (Rose 2014). The EFA GMR (2013) reduces "relevant" in "relevant learning outcomes" to setting proficiency benchmarks that are age and grade appropriate. They view the purpose of monitoring learning outcomes indicators as addressing marginalization and hence recommend disaggregating data according to gender, rural/urban location, and richest and poorest income quintiles.

The Commonwealth Education Ministers proposed a target for learning outcomes that adopted a national level methodology for monitoring relevant learning outcomes through national assessments. National examinations that have a selective function give information on the opportunities open to learners at the end of the basic education cycle and, hence, who benefits from education. They are less useful for cross-national comparison of levels of learning achievement. Indicators that use national assessments, however, need to be supplemented by system level process indicators that set expectations for rigorous and fair assessment and for curricula to be socio-culturally and socio-economically relevant to diverse learners as well as appropriate to age and grade in terms of cognitive demand.

4b. Learning outcomes as indicators of quality

Whilst learning outcomes may be powerful indicators of equity, they can only be partial indicators of other aspects of quality because they tell us little about educational processes. Indeed, national examinations or assessments that are high stakes for students and/or teachers can have considerable washback effects on quality. Some unintended consequences can be avoided by using sample surveys, rather than national examinations and tests but these are no less neutral. The increasing political influence of Large Scale Education Assessments (LSEAs), particularly OECD programmes, has made them complicit in global education governance contributing to a trend of "policy making by numbers" (Sellar & Lingard 2013). Though national representatives participate in planning and designing OECD programmes, non-OECD countries have much more limited participation (Bloem in press). They can, however, assert greater influence over regional LSEAs (e.g. PASEC, LLECE) and the PISA for Development project. Hybrid assessments, such as ASER, Uwezo or Early Grade Reading Assessments (EGRA), are less technically complex but less robust for cross-national comparison. Their simplicity enhances communicability and hence, impact. Whilst they are effective in highlighting where there are serious issues of poor quality and inequity in primary education, they provide limited insight into underlying causes.

International LSEAs are valuable for identifying inequalities between countries but there are diminishing returns in extending them into less readily measurable domains of learning, such as knowledge and skills for sustainable development. This calls for investing in improving the robustness of assessment at the national level and the design of curricula that define the competencies against which students are assessed.

4c. Indicators for equitable outcomes and relevant learning

Figure 5 (also in Appendix 1) suggests a set of indicators for relevant learning that include outcomes indicators of equity, system level process indicators for equity and relevance, and learner level input indicators for readiness to learn.

5. Indicators for teachers and educational processes

Two different genres of literature have highlighted the importance of teachers to education quality – reports and research written or commissioned by rights-based advocacy organizations and large-scale comparative studies, oriented towards informing policy.

5a. Teachers within the rights-based tradition

Proposals for a target on teachers as part of an education SDG address educational processes, which are central to rights-based quality frameworks. Suggestions that have been made for indicators, however, treat teachers as an input focusing on numbers of qualified teachers and their training. Using the term professional development rather than training moves away from top-down implementation of quality improvement to

open up pathways that capitalize and build on the knowledge and expertise within the body of the teaching profession. Professional development is ongoing and it can be more or less formal, ranging from collegial interactions and mentoring relationships to university degrees. Professional associations, for example of subject specialist teachers, bring together individuals with the greatest enthusiasm for extending teachers' knowledge and expertise; these stakeholders can be agentic in developing a vision for quality and disseminating it through the teaching population.

Some research by rights-based organizations has looked at the living and working conditions of teachers, highlighting these as human rights issues in themselves, which also have profound implications for educational quality (VSO 2002; Marphatia *et al.* 2007).

5b. Teachers within policy research

Since the beginning of the 2000s, the teaching profession has become a focus for global policy debate. OECD with the programme Teaching and Learning International Survey (TALIS), the World Bank with SABER-Teachers, and UNESCO are the main policy actors on the international level. In relation to the questions that we should ask of indicators of quality, TALIS in particular is a major research exercise that has much to offer. However, how well TALIS is compatible with a rights-based view of education remains a question.

Regional LSEAs can also and in some cases do collect data from teachers. However, associations between teaching processes and learning outcomes tend to be elusive. Fine-grained qualitative research on teachers shows that the claims teachers make about their practice can diverge, sometimes sharply, from observed practices (e.g. Osborn *et al.* 2000; Schweisfurth 2002). Teachers' responses to questionnaires, therefore, should be regarded as a proxy indicator that gives only partial information on quality. Despite their growing sophistication and yield of large complex data sets, large-scale surveys remain blunt instruments for researching teaching and learning processes.

5c. Indicators for professional teachers and quality teaching processes

It is neither feasible nor desirable to monitor classroom processes at the global level. This is properly the work of the system level, carried out through its inspectorate and other forms of school supervision and evaluation, such as school self-evaluation (Carlson 2009). We thus suggest matching a quantitative input indicator that sets expectations for the frequency of contacts between a school and its supervisors together with a qualitative process indicator that sets expectations for the criteria used to evaluate schools. The latter would need to be elaborated at the national level. The effectiveness of school inspections or evaluations to improve quality depends on clear communication of findings to teachers and the communities served by schools. A third related indicator therefore could address how information on school quality is disseminated and used.

Many countries already have functioning supervision systems and opportunities for

professional development that have been developed on the basis of extensive professional expertise over decades. Indicators need to be more than accountability tools for evaluating education systems but also used to start conversations between and within countries for sharing ideas and professional knowledge.

Further suggested indicators for professional teachers and educational processes are given in figure 7 (also in Appendix 1). They take the input indicators suggested by GCE for number of teachers and their engagement in professional development, rather than training, as a starting point and expand on the inputs, system level processes and school level processes that enable professional teachers and ensure quality schools.

6. CONCLUSION: RE-VISIONING GLOBAL MONITORING

Development goals, targets and indicators, which extend the use of technologies of quantification as tools of global education governance, run the risk of closing off possibilities for education quality. A key question for us is whether it is possible to formulate indicators for the OWG and Muscat targets that transform global monitoring into a process that supports and enables stakeholders at all levels to bring educational processes closer to the ideals of human rights and sustainable development. Believing that education processes and outcomes are situated, we have tried to formulate indicators that encourage problem solving within and across education systems.

The indicators we suggest for relevant learning (see Appendix 1) include quantitative outcomes indicators but also qualitative process indicators intended to stimulate debate, which will lead to the formulation of national and sub-national indicators that support and enable equitable relevant learning for sustainable development. The indicators for teachers (see Appendix 1) draw on the notion of teacher professionalism to challenge treatment of teachers as inputs, passively moulded by training. Instead, they are viewed as active collaborators in creating and implementing a broad and bold agenda that is responsive to the needs and contexts of their pupils and students. Just as education systems should create the conditions for teachers as professionals to be active in creating as well as implementing a vision for education for sustainable development, so global monitoring should create conditions that support re-visioning of education within education systems.

1. INTRODUCTION

1. The goal of Universal Primary Education, first put forward by UNESCO in 1960, was the first of around 50 development goals that followed over the subsequent 40 years (Jolly *et al.* 2005). Since then, remarkable progress has been made in expanding access to primary school to around 90 percent of children at a time of rapid population growth. Now, for the first time, the UN appears be poised to set a goal for education that looks beyond access to target the outcomes of schooling. At the time of this writing, the most significant proposals are for an education and lifelong learning goal that emphasizes quality as well as access:

> Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.
> (Open Working Group for Sustainable Development Goals 2014)

Yet, looking back over 25 years of a global Education for All (EFA) agenda, it is clear that the conceptualization of quality has narrowed from a broad vision that sought to consider all aspects of an education system to a narrow focus on learning outcomes.

2. This paper sets out to reclaim rights-based thinking on quality and apply it to the contemporary task of identifying indicators for a post-2015 education Sustainable Development Goal (SDG). The indicators proposed in this paper offer an interpretation of targets that clarifies a rights-based view of equity and quality in education. The suggested indicators are intended to contribute to a broad and bold agenda for education quality that is compatible with the targets set out in the Muscat Agreement (Global Education For All Meeting 2014) and Open Working Group (OWG) proposal (Open Working Group for Sustainable Development Goals 2014). Much debate on post-2015 targets has assumed the logic of results-based management. In education, this means

that attention has focused on learning outcomes and how to measure them (Center for Universal Education at Brookings 2011; Learning Metrics Task Force 2013; UNICEF/UNESCO 2013). Within a rights-based approach, however, quality is about more than access and outcomes. It inheres in processes that respect and promote children's rights within education. This brings into focus the system level structures that shape processes and enable learning as well as the environment of schools and classrooms and what happens in these spaces. So whilst, monitoring learning outcomes is a significant extension of the last MDG's focus on enrolments, it is not sufficient for realizing the right to education for all children.

3. Setting out a bold and broad agenda for quality at the global level is not without difficulty. It risks subsuming to global level decision-making on policy that is best conducted at the national level and, in so doing, closing off public and professional debate at national and sub-national levels. Global agendas have repeatedly been critiqued for imposing one-size fits all solutions on diverse education systems that have evolved for diverse socio-cultural and political contexts. Hence, we have attempted to formulate indicators that can be adapted and elaborated at national and sub-national levels (Ahmed 2014).

4. Much of the discussion in this paper focuses on developing qualitative indicators for processes. This is because there has already been extensive work undertaken on quantitative measures of learning outcomes and, to a lesser degree, inputs such as numbers of qualified teachers (EFA Global Monitoring Report (GMR) 2013; EFA Steering Committee TAG on the post-2015 indicators 2014; GCE 2014). However, the paper does engage with debate on indicators for learning outcomes as these can have profound implications for learning processes (Goldstein 2004) and governance of education (Dahler-Larsen 2012; Barrett 2013b).

5. The targets addressed in the paper are the "relevant learning outcomes" component of the basic education target and the target for qualified teachers. These two targets need also to be interpreted with reference to a third target, "knowledge, skills, values and attitudes to establish sustainable and peaceful societies" (Global Education For All Meeting 2014: 3), referred to in this document as the Education for Sustainable Development (ESD) target.

6. The paper is organized into two parts. The first part looks back at experience with EFA goals and targets and understandings of quality within the EFA movement to arrive at a framework for formulating indicators. Suggestions for post-2015 indicators are set out in the second part of the paper and are backed by discussion that reviews indicators that have already been developed at the international level and work underway on post-2015 indicators. The paper concludes by arguing that indicators should be used not just to monitor a pre-determined agenda but to support stakeholders across all levels to create and implement a broad and bold agenda for education.

PART I: LOOKING BACK: INDICATORS AND QUALITY WITHIN EFA

7. Over fifty years of setting development goals and 25 years of pursuing EFA at the global level have allowed the development of insights into the role of indicators in catalyzing progress towards achieving targets. Section 2 looks back at experience with the education MDG and EFA goals to ascertain which indicators and targets had greater traction and why, as a basis for identifying a set of criteria for impactful indicators. Section 3 reviews the shifting conceptualizations of quality within EFA that culminated between 2002 and 2006 with the construction of conceptual frameworks. Little work, however, was done towards formulating targets and indicators that would turn these into evaluative frameworks. A framework put forward by Pigozzi (2008) is reviewed and ways in which the formulation of indicators post-2015 could build on this work are identified. Particular attention is given to the analytical distinction Pigozzi makes between the level of the learner and the level of the system. In order to assess the possible unintended consequences of system indicators, we overview programmes conducted by the Organization for Economic Cooperation and Development (OECD) to develop indicators for cross-national comparison. A key concern here is whether system level indicators would close off the space for national adaptation and elaboration of indicators. Section 3 concludes with a framework for designing indicators.

2. THE ROLE OF INDICATORS IN REALIZING TARGETS AND GOALS

2a. Development goals and their influence

8. Development goals can serve four functions (Langford 2012). First, they achieve or reflect a consensus. Second, they prioritize or boost certain areas of action. Langford argues this to be the greatest benefit of development goals. Fukuda-Parr *et al.* (2013), in their critical review of MDG targets for human development and human rights, confirm that the MDGs did shape national policy-making (although not evenly. Third, they can act as a planning tool for policy. However, Langford suggests that this requires the inclusion of conduct-oriented targets, i.e. targets that set out steps or processes to be taken towards achieving a goal. Lastly, they can act as "a soft form of accountability" (*Ibid.:* 12). The MDGs were formulated at a time when audit was increasingly being used as a form of accountability in Northern democracies (O'Neill 2002). By more rigorously defining the aims of international development, the MDGs extended this pattern of results-based management into the international arena (Unterhalter 2005), where it has intensified over the lifetime of the MDGs in part due to the use of evolving technologies of quantification, such as international learning surveys. Languille (2014) provides an example of this in her analysis of how donor partners have attempted to manage education quality in Tanzania by giving increased weight to progress on agreed targets and results in the evaluative framework used to negotiate budgetary support and making extensive use of national examination results as indicators of quality.

9. With respect to education, Verger *et al.* (2012) have argued that together with other international instruments, EFA goals and the education MDG have had such influence internationally that they have come to constitute global education policy. This influence is achieved not just through national policy mechanisms but through the joined advocacy efforts of globally connected civil society organizations (Verger & Novelli 2012). An important part of their argument is that what is often perceived as global or international policy space more properly represents the international influence of agendas generated in geographically specific metropolitan centres (see also Elgert & Krueger 2012). This trend is evident in the post-2015 debate despite the remarkably wide-ranging and long-running process of consultation (Barrett 2013b). Whilst people around the world have contributed to the consultations, the consultations have mainly been coordinated by organizations based in cities such as New York, Washington DC, Paris, London and Montreal, which have then gone on to synthesize and publish the output documents (e.g. Commonwealth Ministerial Working Group on the Post-2015 Development Framework for Education 2012; Learning Metrics Task Force 2013; UNICEF/UNESCO 2013). This context escalates the challenge of setting an agenda that is, in the words of Manzoor Ahmed (2014: 64), "supportive rather than confining" for countries, that is "the floor rather than the ceiling for adapting, broadening and deepening the goals and indicators" at national and sub-national levels.

2b. What does past experience tell us about the impact of indicators?

10. The education MDG (see figure 1) has had remarkable impact against its single target of access to a complete cycle of primary education. Progress was most dramatic between 2000 and the financial crisis of 2008, during which period the number of out of school children in the world nearly halved from 100 to 58 million. However, analysis within the EFA Global Monitoring Report (GMR) (UNESCO, 2010) shows that increases in enrolments and school life expectancies amongst the most marginalized groups, particularly people living in remote and rural areas, has changed very little, even in countries that have raised national averages significantly. Indeed, Carr-Hill (2012) argues that around 55 million children, including nomads, children affected by conflict and emergencies, street children, and children in unstable or multiple occupancy households, are so marginalized as to be missing altogether from these statistics. With the benefits of hindsight, the targets and indicators associated with the education MDG and the EFA goals created an incentive for governments to target "low hanging fruit" by expanding access to primary education for large numbers of children, who are easier to reach in urban or less remote locations (EFA GMR 2013).

11. The 2010 EFA GMR (UNESCO 2010) laid the ground work for equity within the post-2015 agenda by showing how quantitative data can be disaggregated by region/district, rural/urban, gender and income, and intersections between these, to identify groups at the extremes of educational marginalization. The understanding of "marginalization" constructed within this report has informed the Muscat targets, the first four of which are appended with the phrase, "with particular attention to gender equality [or girls and

women] and the most marginalized" (Global Education for All Meeting 2014: 3). The EFA GMR team went on to suggest that post-2015 indicators such as enrolment rates or literacy rates be disaggregated to show the gap and ratio between male and female, urban and rural, the richest and poorest quintile. This would allow monitoring to identify and follow the progress of the most marginalized group. The strategy for addressing equity through identifying and paying particular attention to the most marginalized can also be extended to qualitative indicators (see figure 3).

FIGURE 1

The Education Millennium Development Goal

GOAL 2: Achieve Universal Primary Education

TARGET: Ensure that by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling,

INDICATORS:

1. Net enrollment ratio in primary education
2. Proportion of pupils starting grade 1 who reach grade 5
3. Literacy rate of 15-24 year=olds

12. Whilst enrolments in primary school have expanded, many students, particularly from disadvantaged groups, appear to be benefitting very little from schooling. Surveys assessing learning show that large proportions of children are not achieving even foundational levels of literacy and numeracy (UNESCO 2014). These findings have driven the current impetus to introduce learning targets in the next education development goal (Center for Universal Education at Brookings 2011; UNICEF/UNESCO 2013). Even before poor learning outcomes became an international issue, small scale qualitative research suggested that rapid expansion could lead to poor quality education, with many children in sub-Saharan countries being taught in class sizes of over a hundred in barely resourced schools (Davidson 2004; Croft 2006), sometimes by unqualified teachers (Tanaka 2010). Statistical analysis by the Consortium for Research on Enrolments, Access and Transitions in Education (CREATE) found that many children enrolled in school were attending irregularly, vulnerable to drop-out and effectively excluded from learning (Lewin 2009).

13. Part of the reason for "the learning crisis" (UNESCO 2013) lay in the relative attention given to the three indicators associated with the education MDG. Statistically robust data was available for the first two indicators, Net Enrolment Ratio in primary education and proportion of pupils starting Grade 1 who reach Grade 5 (see figure 1). However, the third indicator of literacy rate amongst 15 to 24-year-olds, which related to schooling outcomes, was harder to measure and consequently given much less attention. Its neglect was inadvertently reinforced by the segmentation of literacy and quality into separate EFA goals. Hence, within the EFA Global Monitoring Report (GMR), progress on literacy rates amongst young adults was reported separately from progress on access and completion of primary schooling. Langford (2012: 8) refers to the prioritization of expansion as an

instance of an indicator reshaping its parent norm as "the focus on access to primary education may have weakened the authority [of] the treaty rights of children to free primary education of a certain quality." A clear lesson is that indicators that are only partially fit-for-purpose in terms of how well they capture the meaning of the target can come to displace or redefine the target.

14. Statistically robust indicators that measure what is quantifiable are very often only partially fit-for-purpose in that they do not capture the full meaning of the target. Therefore, some measurement experts, including Langford (2012), have suggested using them in combination with less robust but more fit-for-purpose indicators. For example, Unterhalter (2014) proposes that for the education goal, the proportion of children enrolling in Grade 1 successfully completing an examination at the end of their basic education cycle, a relatively robust indicator for which verifiable data is available, could be used alongside a new yet to be developed indicator gauging engagement with lifelong learning, knowledge about sustainable development, equity, and attitudes towards violence against women, which would be derived from survey data. Her proposal for the second indicator conforms to Langford's (2012: 20) suggestion that not all post-2015 indicators may have comparable and robust data available at the beginning of the 2015-2030 period but there should be a realistic prospect for developing measures during the SDGs' lifetime.

15. Nonetheless, experience with the education MDG suggests that the intended rebalancing effect of new, more complex, less robust indicators is likely to be limited because they tend to gain less traction. Ensuring that the whole set of indicators associated with a target is fit-for-purpose will depend on monitoring and reporting mechanisms that amplify the less statistically robust indicators. This may be done through, for example, discursive reporting that looks across available evidence against all indicators to construct a broad picture of progress against the target. At the global level, it also means that continuing to improve precision of statistical measures for which relatively robust data is already available (such as enrolment, transition to Grade 5 or measures of literacy skills) will produce diminishing returns in terms of supporting balanced progress against targets. Instead, resources should be invested in developing indicators and collecting data for the aspects of targets that are complex and harder to measure, such as "knowledge, skills, values and attitudes for sustainable development," in order to achieve a more complete picture of progress. This will also involve supporting the development of indicators at the national and sometimes sub-national level, as some qualitative terms such as "skills for decent work and life" can only be interpreted fully with reference to specific contexts.

16. Indicators are more likely to have traction if they are comprehensible and valued by implementers at various levels, including educational professionals and civil society activists, and can harness the support of wider society. Langford (2012: 20) calls this criteria "communicability" or "salience." Unterhalter (2014) argues this point with reference to findings of the Gender, Education and Global Poverty Reduction Initiatives

(GEGPRI) project, which examined the implementation of MDG 1, 2 and 3 in Kenya and South Africa. She argues that the power of the development goals to promote equity depends on education professionals at different levels and the broader public being able to understand and engage in the review of indicators and targets. This relates to the broader question of participation in monitoring that applies to all levels from the national downwards. Public debate and participation in decision making are fundamental to social justice (Fraser 2008; Sen 2009). Within rights-based frameworks for conceptualizing education quality (see section 3a) participation of stakeholders in decision-making appears as a central dimension of quality.

2c. What version of quality should indicators promote?

17. Langford's critique of the education MDG is also a reminder that the fitness of purpose of indicators needs to be evaluated not only with reference to the agreed wording of the target but keeping in mind the human rights that the target aims to realize. This, however, is far from simple for education because the right to education is not just about rights to outcomes but rights to processes (McCowan 2013). This has also been expressed by referring to the rights *to* education (access), *in* education (quality, processes) and *through* education (outcomes) (Subrahmanian 2002). The proposals for a goal capture this by referring to "quality education." It is the targets and their indicators, though, that elaborate what "quality" means in the context of a global development framework.

18. The education development goal is just one of 17 goals proposed by the OWG that together constitute an agenda for sustainable development. Indicators need to interpret the targets in ways that are consistent with the overarching purpose of promoting development that is economically, socially and environmentally sustainable (Barrett 2013a). Another way of expressing this is that the indicators need to dovetail and not conflict with targets and indicators for other SDGs, such as the goal for healthy lives and well-being, empowering women and girls, decent work for all or promoting peaceful and inclusive societies for sustainable development. The Muscat and OWG proposals make explicit reference to sustainable development in the ESD. ESD is a complex field but consistent with a rights-based view of quality in its insistence on continuity between processes and outcomes (Tilbury & Fien 2009; McCowan 2013). Problem-solving and creativity are essential to both. There are differences, however. The environment is foregrounded in ESD, which recognizes human dependence on the environment, but absent from the human rights based understandings of development. Human rights are written in the present tense. In education, they generate an imperative to fulfill entitlements for today's children, young people and illiterate adults. Sustainable development, by contrast, has a future tense and makes the wellbeing of future generations as central to decision making as that of today's children.

2d. Criteria for indicators that create and sustain a broad agenda for quality

19. This discussion of the experience with the education MDG and EFA can be distilled into six criteria for indicators, which need to be met if they are to set an implementable agenda for quality and equity in education.

1. **Compatible with the right to education:** Indicators together promote the rights of children to a *quality* education as set out in the Convention on the Rights of the Child and other relevant UN Conventions.
2. **Contribute to sustainable development:** Indicators articulate with other SDGs to promote economically, socially and environmentally sustainable development.
3. **Relevant to diverse contexts:** Indicators support national policy-making and relevance for diverse country contexts by allowing for adaptation and elaboration at the national level.
4. **Equity:** Indicators differentiate between more and less disadvantaged groups so that the most marginalized are recognized and given particular attention.
5. **Fit-for-purpose:** The whole set of indicators associated with a target should together promote all aspects and dimensions of the target.
6. **Communicable and salient:** Indicators are readily comprehensible and meaningful to implementers, including education professionals, EFA activists and wider society.

20. The first three of these criteria require that education targets have qualitative indicators for processes. They also necessitate developing new measures, particularly at the national level. Creating and sustaining a broad agenda for quality will be dependent on all indicators receiving attention. There must, therefore, be a realistic possibility of developing measures for new indicators within the next few years. It is also important to find ways to amplify qualitative indicators through advocacy and mechanisms of reporting that prevent the agenda snapping back to one narrowly focused on what is most readily measurable.

3. CONCEPTUALIZING AND TARGETING QUALITY IN EFA

21. The post-2015 proposals for education (Global Education for All Meeting 2014; Open Working Group for Sustainable Development Goals 2015) look much more like the current EFA goals than the education MDG. This section, therefore, looks back at the history of how quality has been conceptualized in the EFA movement from 1990 onwards as a source of ideas for targeting quality post-2015. Quality is targeted in Muscat and the OWG proposal in two ways. First, the target for each educational level is specified in terms of the outcomes, knowledge, attitudes and/or skills to be achieved. For the basic education cycle, defined in Muscat to include lower secondary as well as primary, the target is not just completion but to achieve "relevant learning outcomes." Both documents propose a target for increasing the supply of qualified teachers. OWG also includes a target for the built school environment.

3a. Rights-based understandings of quality

Learning for All in the World Declaration on Education for All

22. In some respects the proposals for a post-2015 target for basic education represent a return to the spirit of the 1990 World Declaration on Education for All, adopted at Jomtien. This emphasized learning rather than schooling and a metrics of learning achievement to the extent that 25 years on, sections of the World Declaration seem strikingly contemporary:

> Whether or not expanded educational opportunities will translate into meaningful development – for an individual or for society – depends ultimately on whether people actually learn as a result of those opportunities. ...The focus of basic education must, therefore, be on actual learning acquisition and outcome, rather than exclusively upon enrolment.... It is, therefore, necessary to define acceptable levels of learning acquisition for educational programmes and to improve and apply systems of assessing learning achievement. (World Conference on Education for All 1990: article 4, p. 5)

23. The World Declaration also foreshadowed the attention to basic and not just primary education in the Muscat Agreement. It called for the universalization of basic education, allowing the basic education cycle to be defined at the national level, and for "broadening the means and scope of basic education" (World Conference on Education for All 1990: article 5, p. 5), including early childhood care and education and formal and non-formal programmes for youth and adults.

The EFA quality goal and the Dakar Framework for Action

24. The perils of rapid expansion were recognized in 2000 and the sixth EFA goal was intended to send out a clear message that expansion has to go hand-in-hand with quality improvement. However, in doing so it conflated quality improvement with the achievement of measurable learning outcomes:

> Improving every aspect of the quality of education, and ensuring their excellence so that

> recognized and measurable learning outcomes are achieved by all, especially in literacy, numeracy and essential life skills. (World Education Forum 2000: 17)

25. The Dakar Framework for Action did not specify indicators for quality. The EFA Global Monitoring Reports have over the years reported on pupil:teacher ratio, the proportion of teachers qualified, learning achievement as indicated by various large-scale education assessments, such as PISA or SACMEQ, and expenditure on education. In later reports, expenditure was defined as a distinct issue and from 2011 allocated a dedicated chapter. Post-2015 proposals include a dedicated finance target. As increasingly detailed information on learning achievement became available, including for children in lower primary, this was given greater page space than teacher indicators. The latest GMR (UNESCO 2014) additionally flagged measures of textbook availability and school infrastructure as other important input indicators. Another change is that in early reports, reporting on goal six was prefaced by cautions regarding the limitations of available indicators, carefully described as "proxies" for quality. As more data became available on trends in learning outcomes, these cautionary notes were dropped, and the GMR asserted its findings with a greater sense of urgency.

26. The limited set of quality indicators available to the GMR team do not fully capture the understanding of quality within the Dakar Framework of Action, which refers to quality in terms of processes – "what takes place in classrooms and other learning environments" – and benefits to the learner:

> A quality education is one that satisfies basic learning needs, and enriches the lives of learners and their overall experience of living. (World Education Forum 2000: 17)

Measurable learning outcomes, therefore, are the indicator, or proxy, for these broadly conceived benefits for the learner.

27. Within Dakar, eight process and input factors are associated with a good quality education:

1. healthy, well-nourished and motivated students;
2. well-trained teachers and active learning techniques;
3. adequate facilities and learning materials;
4. a relevant curriculum that can be taught and learned in a local language and builds upon the knowledge and experience of the teachers and learners;
5. an environment that not only encourages learning but is welcoming, gender-sensitive, healthy and safe;
6. a clear definition and accurate assessment of learning outcomes, including knowledge, skills, attitudes and values;
7. participatory governance and management; and
8. respect for and engagement with local communities and cultures.

(World Education Forum 2000: 17)

Rights-based quality frameworks

28. Between 2000 and 2008, rights-based organizations suggested frameworks for conceptualizing quality based on the principles enshrined in the Convention on the Rights of the Child that replicated some of the eight factors in Dakar. UNICEF's "five key elements that affect the quality of learning" (UNICEF 2008) borrowed from a list published by the GCE (GCE 2002) and orientated it towards education for girls. Outcomes was just one of five principles for judging education quality, the others related to children's readiness for learning, the learning environment, content and processes (see figure 2). The original GCE list also included a sixth dimension, responsiveness, which concerned the relationship between schools and communities expressed as the extent to which education responded to the views and concerns of stakeholders.

29. The 2005 EFA GMR, which focused on quality, proposed a "framework for understanding, monitoring and improving education quality" with a similar structure of five central dimensions: learner characteristics; context; enabling inputs; teaching and learning; and outcomes (UNESCO 2004: 36). Each of these dimensions brought together elements that are substantively different in nature and should be treated analytically in different ways. So, "learner characteristics" included school readiness, which highlights the link between early childhood care and education and aptitude. The "context"

FIGURE 2

Five Key Elements That Affect Education Quality

1. What students bring to learning. What experiences does the learner bring to school, and what particular challenges does she face? Has she been affected by emergencies, abuse, daily labour or AIDS? Has she had a positive, gender-sensitive early childhood experience within her family, her community and her preschool? How different is the language of her home from the language of her school? Has she been sufficiently oriented to the rhythm of schooling?

2. Environment. Is the learning environment healthy, safe, protective, stimulating and gender-sensitive?

3. Content of education. Are the curriculum and materials relevant? Do they impart basic skills, especially in literacy and numeracy? Do they promote life skills and knowledge areas such as gender, health, nutrition, AIDS prevention, peace, or other national and local priorities? How does the content of curriculum and learning materials include or exclude girls?

4. Processes. Are teachers using child-centred teaching approaches? Do their assessments facilitate learning and reduce disparities? Are classrooms and schools well-managed? Are the methods of teaching, learning and support – whether from supervisors, teachers, parents or communities – enhancing or undermining girls' achievement?

5. Outcomes. What outcomes of basic education do we expect for girls? How can we document how well girls are learning and how well the curriculum furthers their future growth? Learning outcomes should be linked to national goals for education and should promote positive participation in society. (UNICEF 2008)

dimension included enabling inputs such as teaching and learning materials, physical infrastructure and human resources and school governance, which could be classified as a process. Teaching and learning overlapped with UNICEF's processes element but included class size, which may also be regarded as an infrastructure issue. Outcomes included learning achieved by individuals (literacy, numeracy, life skills, values, etc.) and social benefits, which depend on relevance of learning and opportunities for individuals to use their learning to the benefit of society. Context is conceptualized as both influencing and being influenced by the other four elements and includes the micro-context of teacher; peer effects; parental support; the systemic context (national standards, national governance and management); and macro socio-economic and socio-cultural context (national standards, labour market, globalization, etc.).

30. The 2005 GMR used its framework to structure its analyses of how various variables result in improved learning outcomes. Its analysis therefore reproduced the sixth goal's conflation of quality and measurable learning outcomes. This conflation is avoided by the UNICEF framework, which asks questions regarding how well education systems respect and promote human rights not just through the learning outcomes they deliver but also through their processes.

31. More recently, the Beyond2015 campaign provided a definition of education quality based on a review of several key documents published by rights-based organizations, including the Delors Report (Delors *et al.* 1996), the Dakar Framework for Action, the 2005 EFA GMR, the work of UNICEF and GCE. The review emphasized cultural and value-based outcomes from education less readily measurable than skills and knowledge:

> Quality education, therefore, builds knowledge, capabilities and life skills and values, and develops the creative, social and emotional capabilities of learners. It fosters broad cognitive and personal development, including critical and higher order thinking, problem-solving, self-discipline, and can support active citizenship, leadership and more. Quality education must also be non-discriminatory; equality is in itself a key component of quality education. (Beyond2015 2013: 11)

32. The authors also noted that consistent attention was paid across the rights-based literature to three systemic elements of a quality education: teachers and teaching; curriculum and content; and the learning environment. It called for qualified, skilled teachers who are knowledgeable with respect to both their subject area and pedagogy, including learning assessment. It stated that a comprehensible, relevant and meaningful curriculum should be inclusive, promote learners' rights, make use of children's mother tongue and include play, sport and creative activities as well as life skills. "Learning environment" in this review referred to school infrastructure and facilities.

33. The discussion so far in this section shows that a quality of education consistent with human rights generates outcomes beyond literacy and numeracy that benefit learners throughout their lives and also benefit the societies in which they live. For individuals,

instrumental benefits include basic learning needs, particularly skills in literacy and numeracy, but stretch much further. A quality education builds a broad range of knowledge, skills and values that encompass the cognitive, social, emotional and creative domains. It develops capabilities for contributing towards national development goals and positive participation in society, including leadership and citizenship skills, knowledge and skills related to gender awareness, health, nutrition, peace and respect for the culture of others. These outcomes are realized through classroom and school processes that are directly experienced by the learner, processes that recognize and respect what the learner brings – her socio-cultural background, identity and prior knowledge; that engage with the learner's community; that ensure the learner is well-nourished and ready to learn; that create a safe and healthy learning environment for girls and boys; and processes that are equitable. These processes are enabled through a series of system level inputs and processes: adequate physical infrastructure; well-trained qualified teachers, relevant curriculum and learning materials, participatory governance and management, and accurate assessment of learning.

34. In short, we already have within the body of human rights literature associated with EFA a broad and bold vision for education quality. The Muscat and OWG proposals expand this agenda through integrating it with the sustainable development agenda. But they also crop the agenda by containing it within targets for enrolment and measurable learning outcomes; less measurable outcomes of knowledge, skills, values and attitudes for ESD and peace/citizenship; and inputs of infrastructure and teachers. Processes do not make the cut.

35. Solidifying a broad and bold agenda through constructing indicators for an expansive vision for quality is a challenging task. Not least because the benefits of quality education listed above are situated within diverse geographies, economies and cultures. Within the context of a global policy field heavily influenced by the logic of results-based management (see section 2a), measures of a limited set of basic learning outcomes, originally regarded as proxies, have come to be widely accepted as an accurate and objective representation of quality. Hence, the broad and bold vision for quality is displaced by a narrow concern with standardized learning assessments.

3b. Defining system level indicators

36. Whilst much work has been done that defines a broad rights-based framing of quality, less work has been done to construct a system of indicators for monitoring quality. Pigozzi (2008), former Director of the Division for Education Quality at UNESCO, went the furthest in seeking to develop a set of indicators for quality derived from a rights-based framework. Her framework revolves around learning as the heart of the educational endeavor but identifies two levels of organization that enable learning. The first level, the level of the learner, closely resembles the UNICEF framework in Figure 2. However, a second level, that of the learning system, is wrapped around this, also with five dimensions:

- Structures management and administration to support learning
- Implements relevant and appropriate policies
- Promotes the establishment of legislation supportive to learning
- Restructures resources for learning
- Measures learning outcomes (Pigozzi 2008)

37. Pigozzi goes on to suggest targets and indicators, although she terms these "indicators" and "measures," respectively. Many of her "measures" indicate what is to be measured without identifying data sources. For both the level of the leaner and the level of the system, indicators are a mix of quantifiable outcome and qualitative conduct indicators. The latter relate to actions or policies taken towards achieving a target. For example, indicators for the dimension "seeks out the learner" are disaggregated Net Enrolment Rates (NER), "special efforts to be gender-responsive" and "affirmative actions in place for the hard to reach" (Pigozzi, 2008: 13). At the system level, taking "implementation of good policies" as an example, the suggested indicators are:

- Education institutions that meet health and safety standards
- Education staff know and employ rules and practices that support [education institutions as workplaces policies] and their professional development
- Student enrolment and achievement data are consistent with population distribution data (Pigozzi, 2008: 14)

38. Some proposals generated by the post-2015 debate include system level conduct indicators. For example, GCE (2014) suggests a finance target indicator: "Development of a fully costed national education plan and a financing strategy" (GCE 2014: 5). The Post-2015 Education Indicators TAG (2014) has suggested include "nine years of free and compulsory basic education in legal/institutional frameworks" as an indicator for the basic education target. However, system level indicators have not been developed systematically across the targets.

39. Langford (2012) notes that conduct indicators can set out an action-oriented agenda that focuses on steps to be taken, rather than a compliance agenda that constantly looks backwards at what has been achieved so far. Hence, they can enhance the contribution development goals make to planning. However, putting in place policies does not ensure their implementation, as seen, for example, with policies prohibiting corporal punishment across a number of sub-Saharan African countries. Langford suggests four circumstances under which a conduct indicator might be used:

- where there is consensus that a particular intervention is a necessary and largely sufficient condition for achieving an outcome
- the target is derived or aligned with international standards or obligations concerned with conduct

- outcome indicators are less robust than conduct indicators
- an outcome indicator can only be interpreted with the use of a conduct indicator (Langford 2012: 24)

40. Pigozzi's index of quality is a long way from providing a full set of indicators that can be used as part of a global monitoring architecture. However, it is significant in taking seriously the view represented in the Dakar Framework for Action that teaching and learning is enabled by education systems and taking a first step in developing indicators for those systems.

41. In summary, three main lessons can be drawn from Pigozzi's work and the rights-based frameworks in general. First, monitoring of quality needs to ask searching questions of the system as well as learner level. Second, devising indicators that address the system level is far from easy. Third, conduct indicators have a role to play in monitoring quality.

Problems with system level indicators

42. OECD programmes have arguably taken the lead in designing system level indicators for the purpose of cross-national comparison. The definition of system level indicators entails considerable challenges, the scale of which is, for example, indicated by the logistical and procedural efforts in OECD programmes to ensure "cross-culturally valid" data (see OECD Technical Reports published by, for example, the PISA and TALIS programmes). Yet, the issue of cultural bias remains one of the most heavily criticized aspects of OECD programmes. The critique suggests that despite the claims made to objectivity, policy recommendations emanating from OECD programmes are based on a normative developmentalism, which disregards local, national and regional diversity in institutional arrangements as well as the fact that education systems are embedded in socio-cultural environments with distinctive traditions, norms and practices (Alexander 2000; Goldstein 2004; Nardi 2008).

43. In recent years, the World Bank has shown considerable interest in education systems (World Bank 2011). Its Systems Approach for Better Education Results (SABER) programme is intended to provide a set of benchmarks for assessing education systems. Initial work on this has focused on benchmarking policies, with recommendations suggesting a single direction of travel. The World Bank's (2011) Education Strategy and SABER have both been critiqued for promoting neoliberal policies in education (Robertson 2012; Robertson *et al.* 2012; Verger *et al.* 2012). Critics point out that the recommendations for decentralisation and liberalisation of the education sector stand in contradiction to the strong centralised planning that characterizes governance in countries such as South Korea, Singapore and Cuba, which have improved equity and quality.

44. The settlement of a global educational policy field in recent decades has also involved the rise of a new group of for-profit policy actors that have issued high-profile reports. In this respect, the McKinsey reports (Barber & Mourshed 2007, Mourshed *et al.* 2010) and Pearson's Learning Curve project (The Economist Intelligence Unit 2014) stand out as amongst the most influential. These reports are prone to critique even more than the OECD programmes due to their endorsement of one-size fits all policy solutions (Morris in press). On this basis, Coffield (2012) argues that the models proposed in the McKinsey reports are unsophisticated, impracticable and undemocratic.

45. The logic that underpins all these approaches to systems benchmarking is still one of results-based management that defines quality in terms of outcomes. Indicators are justified on the basis of their association with improved learning outcomes as measured through performance in standardized tests and ultimately in terms of the assumed association with improved national competitiveness within the global economy. These assumptions are distinct from the logic of a rights-based perspective. As demonstrated above, within human rights perspectives outcomes are just one dimension of quality and processes and inputs have intrinsic value for enacting rights within education, independent of associated outcomes. Indeed, the distinction between processes and outcomes dissolves when outcomes such as attitudes and skills for contributing to peaceful societies are considered.

46. Nonetheless, the critique of benchmarking within the work of OECD and the World Bank and the McKinsey reports has pertinence for the use of system level indicators within a global monitoring framework. It demonstrates that systems level indicators, just like measures of learning outcomes, can work to diminish the agenda for education quality. Indeed, the very definition of indicators as a driver for education reform is likely to have constitutive effects in the realms of culture and political life (Dahler-Larsen 2012). The stakes are raised in this respect when system level indicators are to be negotiated and defined at the international and global level.

47. The sections below will explore the potential of system level indicators to supplement learner level indicators within a rights-based post-2015 agenda that sets a floor and not a ceiling for education quality (Ahmed 2014). However, given the critiques of existing systems level indicators, there is a need to proceed with caution. Candidate indicators should be assiduously assessed according to their potential to support context-sensitive problem solving at the national and local level by allowing for adaptation and a variety of approaches within an overarching rights-based framework. The next section, therefore, sets out some questions to guide the assessment of indicators.

3c. A framework for designing indicators

48. To conclude Part 1, we take the vision for education quality derived from discussion in section 3a together with the distinctions between levels and types of indicators presented in section 3b to start constructing a framework for designing indicators (see

figure 3). At this stage, we do not formulate indicators but simply match up some of the characteristics identified with a quality education within rights-based literature with the different types of indicator discussed. We also suggest whether indicators can be set at the international or national level. Most indicators set at the international level will require adaptation and elaboration for the local level. This also applies to outcomes indicators because the sustainable development benefits of education are by nature situated and cannot be defined with precision at a universal level. The processes through which indicators are developed across levels should be consistent with principles of participatory governance. Participatory formulation of national and sub-national indicators will be easier if international indicators are communicable and salient.

FIGURE 3. A FRAMEWORK FOR DESIGNING INDICATORS

Level of system	Level of learner	Types of indicator *Set at international or national level?*
OUTCOMES		
Curriculum aims to develop knowledge, skills, values and attitudes that allow learners to contribute to other SDGs. e.g. improved nutrition and sustainable agriculture; empowerment of women and girls; peaceful and inclusive societies; sustainable economic growth; conservation of ecosystems; combat climate change	Knowledge and skills for participating in civil society – literacy, numeracy, leadership, knowledge skills, values and attitudes for responsible citizenship, including emotional skills Economic domain – knowledge and skills for employment and productive work that contributes to sustainable development. Cultural domain – participating in artistic, cultural (including sports) and intellectual life of society. Environmental domain – knowledge, skills, values and attitudes for environmental conservation and restoration.	conduct – policy *International, elaborate at national level* outcomes – quantitative *Mainly at regional or national level because benefits of education are situated. Aggregated monitoring of some limited learning outcomes may be possible, e.g. literacy and numeracy, attitudes.* processes – qualitative (because some outcomes are continuous with processes) *National or sub-national, as benefits related to SD are situated*
PROCESSES		
Participatory governance Accurate assessment of learning	Learners are enrolled into appropriate education programme Recognize and respect what the learner brings Engage with the learner's community Ensure learner is well-nourished and ready to learn Maintain safe and healthy learning environments for girls and boys Formative assessment Appropriate pedagogies Participatory school management	conduct – policy *International, elaborate at national level* processes – qualitative *International, elaborate at national level where it is possible to more precisely define aspirations for processes*
INPUTS		
Trained, qualified teachers	Numbers of teachers Infrastructure of schools Relevant learning materials Free school meals as needed Clean drinking water and sanitation	Processes - qualitative (e.g. quality of teacher education) *International, elaborate at national level* Input – quantitative *International level. Will need to be adapted according to capacity to collect and analyze data at national level and baseline from which a country is starting.*

PART II: LOOKING FORWARD: RIGHTS-BASED INDICATORS FOR THE EDUCATION SDG

49. The second part of this paper focuses on indicators for two of the post-2015 targets. Section 4 focuses on indicators for relevant learning outcomes within the basic education target. It discusses how "relevant" is interpreted and overviews the kinds of learning assessments currently being used to monitor learning outcomes globally. An argument is then put forward for supplementing learning outcomes indicators with qualitative indicators at the system and learner levels that direct action towards creating necessary conditions for learning. Section 5 turns to indicators for the teachers' target showing how previous work has tended to construct teachers as inputs for quality, moulded through training, rather than professionals, who are agentic in improving quality. The *Teaching and Learning International Survey* (TALIS) conducted by the OECD is briefly overviewed to suggest what may be possible for monitoring and learning. The section ends by arguing that input indicators for teachers should be supplemented by qualitative system level indicators focused on how teachers are supervised and how their work is assessed as well as enabling conditions for teacher professionalism.

4. INDICATORS FOR "RELEVANT LEARNING OUTCOMES"

50. The Muscat and OWG targets are concerned with learning outcomes of various kinds (see figure 4). In this paper we focus on relevant learning outcomes for the basic education target but interpret relevance with reference to the other targets that speak to educational outcomes. We look at some suggestions for indicators for learning outcomes, with a view to ascertaining how relevance is implicitly interpreted.

51. Our understanding of relevance in education is underpinned by previous work on education quality and social justice (Tikly & Barrett 2011; Barrett 2011). We view relevance as having socio-cultural and socio-economic dimensions concerning preparation for life

FIGURE 4. LEARNING OUTCOMES IN THE MUSCAT AGREEMENT TARGETS

Target	Learning Outcomes
Basic Education	relevant learning outcomes
Youth and adult literacy and numeracy	Proficiency in literacy and numeracy sufficient to fully participate in society
ESD	knowledge, skills, values and attitudes to establish sustainable and peaceful societies"
Life skills	knowledge and skills for decent work and life

in local and national contexts and participating in the benefits of globalization. Socio-cultural meaningfulness demands recognition of learners' diverse and multiple socio-cultural identities. Socio-economic relevance concerns how well education prepares learners for sustainable livelihoods. This view of relevance is consistent with the ESD target and skills for work and life target.

4a. Learning outcomes as indicators of equity

52. Relevance as defined above, however, has not been addressed by much of the existing work on post-2015 indicators. Measures of learning are commonly treated as unproblematic indicators of quality (e.g. Filmer *et al.* 2006) and important indicators of equity, allowing us to identify who is excluded from a quality education (Rose 2014). The most recent document produced by the Technical Advisory Group (TAG) to the EFA Steering Committee, charged with developing indicators for the Muscat targets, addresses the four outcome targets through quantitative output measures of participation in education and outcome indicators of learning. The free and compulsory component of the basic education target is addressed through a conduct indicator. For ESD, the document states targets that currently do not exist but that citizenship outcomes could potentially be measured through household survey and learning assessments.[1] The learning outcomes indicators suggested for the basic education target in this and earlier documents published by UNESCO (EFA GMR 2013) are measures of proficiency in literacy and numeracy. They clarify that the "relevant" in "relevant learning outcomes" refers to setting proficiency benchmarks that are age and grade appropriate. The purpose of monitoring learning outcomes indicators is taken to be to promote equity by focusing attention on groups marginalized through education, in other words, groups that may be enrolled in school but are not learning (UNESCO 2010; EFA GMR 2013; Rose 2015). The EFA GMR (2013) team took a similar position, recommending that equity be addressed through setting benchmarks for the gap between and ratio of proficiency levels for males and females, rural and urban learners, and richest and poorest quintile by income.

53. The proposal from the Commonwealth Education Ministers is of interest because it envisions a form of learning indicator that covers a broad range of competencies:

> Successful achievement of national learning outcomes in cognitive, affective and psychomotor domains of both primary and lower secondary cycles at age appropriate levels up to the age of 15 years. (Commonwealth Ministerial Working Group on the post-2015 development framework for education 2012: 6)

Thus, it adopts a national level methodology for defining and monitoring learning outcomes. In so doing, the Commonwealth Education Ministers illustrate how "relevant

1. Precedence exists for international comparison of citizenship knowledge and understanding. The International Association for the Evaluation of Educational Achievement (IEA) is preparing for its second International Civics and Citizenship Study (ICCS) in 2016, targeting Grade 8 students and teachers. ICCS develops region specific modules for Asia, Europe and Latin America that supplement its international modules (IEAc. 2013).

learning outcomes" may be interpreted at the national level through reference to existing curricular objectives. This allows for contextual specificity and communicability at national and sub-national levels.

54. National assessments may take the form of surveys that assess a sample of learners from a grade or age cohort, or national examinations and tests that are taken by a whole cohort. National examinations that have a selective function play a key role in determining future educational and employment opportunities. Hence, they offer information not just on what learners can demonstrate they have learned under test conditions, but the opportunities open to them after they complete basic education. Scholars adopting a social justice perspective, particularly those using Sen's capability approach, argue that benefits and opportunities are the appropriate space for evaluating equity (Unterhalter & Brighouse 2007; McCowan 2010; Tikly & Barrett 2011; Barrett 2011a). Furthermore, data can readily be disaggregated by gender, age of candidates and the location of schools to identify patterns of marginalization. It is valuable therefore to include examinations data in national level monitoring in education but with indicators disaggregated in the ways suggested by the EFA GMR (2013) to give information on equity.

55. National examinations are less useful for cross-national comparison of levels of learning achievement since different countries assess students against different, although overlapping, sets of competencies. Further, benchmarks that determine what counts as a pass or fail may be influenced by factors extrinsic to basic education, most especially the availability of post-secondary education, which in turn may be influenced by the capacity of the formal economy to absorb graduates from post-secondary levels of formal education (Palmer *et al.* 2007).

56. The Commonwealth Education Ministers proposal assumes the sufficiency of nationally defined learning outcomes in terms of being relevant to the national context and appropriate in cognitive demand. It also takes for granted the robustness of national assessments and that they are rigorous and fair. In practice, both of these are problematic in many under resourced education systems (Pritchett & Beatty 2015). Where national assessment data is used in outcomes indicators, attendant system level process indicators are needed that set expectations for the quality of assessment and curricula. These need to be designed in ways that do not compromise contextual socio-economic and socio-cultural relevance of curricula by over-specifying expectations or limiting possibilities. Their implementation should be directed towards expanding the capacity of curriculum authorities and examination councils to design curricula and assessment tailored to their own country context. As teachers' work is usually evaluated in terms of their pupils' or students' performance in national examinations and the conformity of their teaching with the curriculum, such system level process indicators are likely to have a more direct influence on processes of teaching and learning than LSEAs. They should not, however, displace measurement of equity in learning outcomes.

57. To conclude, national examinations data, when disaggregated according to student and school characteristics, can be used to indicate equity along with data on who progresses to what kind of post-basic education or employment. They need, however, to be supplemented by system level process indicators that set expectations that assessment is rigorous and fair, and that curricula are relevant to diverse learners and appropriate to age and grade in cognitive demand. Further, data from examinations should never be treated as a proxy for all aspects of quality.

4b. Learning outcomes as indicators of quality

58. Whilst learning outcomes may be powerful indicators of equity, they can only be partial indicators of other aspects of quality. Performance in standardized pen and paper tests tell us little about educational processes. Indeed, national examinations or assessments used by governments to monitor school performance that are high stakes for students and teachers can have considerable washback effects on quality. For example, learning outcomes in academic subjects such as literacy, mathematics and science may be raised through the use of restricted pedagogies that focus on "teaching for the test" at the cost of time spent on other parts of the curriculum, including those that develop creative and problem-solving skills. Competition to succeed in national examinations can drive up learning outcomes but also fuel shadow education systems that exacerbate socio-economic inequalities and in some instances deny children their right to play and leisure (Bray 2007; Sobhy 2012). The effects of high stakes testing may be different for schools serving large numbers of disadvantaged learners, where concerns about underperformance create pressures to focus on literacy and numeracy at the expense of critical thinking, argumentation and abstract thinking skills developed through, for example, the study of literature or logic (Anagnostopoulos 2007). High stakes assessment can also work to limit teaching and learning time spent on unexamined areas of the curriculum, which often include life skills and citizenship. The drive to improve learning outcomes has also become a justification for low-fee private schools, which overlooks the right to free quality primary education (Robertson *et al.* 2012; Srivastava 2013).

59. Some of these potential unintended consequences can be avoided by using learning surveys, rather than national examinations and tests. The EFA GMR (2013) team proposed the use of cross-national large-scale educational assessments (LSEAs) including OECD and International Association for the Evaluation of Educational Achievement (IEA) studies used in conjunction with regional LSEAs such as LLECE, SACMEQ and PASEC. These studies have the advantage of being internationally recognized statistically robust measures, large enough to draw on world leading technical expertise primarily located in North America, Australia and Europe. As well as assessing learning outcomes through standardized tests, LSEAs collect information on learners' home background, school characteristics, classroom practices and system level curriculum and policy so that associations can be explored. Hence, they have played a critical role in highlighting inequalities in learning outcomes between high and low income countries and between different socio-economic groups within countries (Bloem 2013).

60. International LSEAs, however, are no less neutral as measures of education quality than national assessments. As noted in section 3b, OECD programmes and IEA studies have been criticized for cultural bias. Neither are they politically neutral. The increasing political influence of international LSEAs, particularly OECD programmes, has made them part of the architecture of global education governance. This is most visible with respect to extensive reference made to OECD's PISA study in policy debate in European Union and other OECD countries, which researchers have characterized as policy making by numbers (Grek 2009; Sellar & Lingard 2013). Ironically, the apparent objectivity of rigorously and technically highly complex learning surveys lends them political potency for actors seeking to use their authoritative findings to justify ideologically motivated policy decisions (Chung in press). For example, in the first decade of the century, PISA results were used to justify very different directions of change in education policy in Sweden and Germany (Ringarp *et al.* 2010).

61. The OECD has been observed to actively seek to extend its influence within educational governance in Europe on the basis of LSEAs (Grek 2009). Publications from these programmes and the communication activities of its senior officers are characterized by a neo-liberal discourse that emphasizes education's role within a competitive global skills market. However, OECD is not a monolith and the decision-making committees shaping PISA and other OECD programmes are sites of contestation between the national interests that make up the organization. Business interests and teacher unions are also involved in negotiating OECD programme designs, although their role as observers in the relevant bodies renders them less influential.

62. Non-OECD countries have less opportunity to influence the design of OECD programmes, which can limit the responsiveness of PISA to pressing policy issues in these countries (Bloem in press). However, they can assert greater influence over regional LSEAs and the PISA for Development project. Regional LSEAs, therefore, tend to be more responsive to national curriculum objectives and adapted to probe inequalities of greatest concern. For example, in response to the priorities of participating ministries, the 2007 SACMEQ study included a HIV and AIDS Knowledge Test and collected data on students' disabilities. SACMEQ and PISA for Development also have an explicit objective of developing capacity within ministries of education to administer surveys and analyze data. One drawback of many LSEAs is that they are administered by ministries of education to grade cohorts through schools and hence do not compare across children and young people who are in and out of school. Their technical complexity may diminish communicability, particularly for independent advocacy groups.

63. Hybrid assessments, such as ASER, *Uwezo* or Early Grade Reading Assessments (EGRA), are less technically complex but also less robust, particularly for the purpose of cross-national comparison. They have been designed to target lower primary learners. Wagner (2010: 747) describes them as "smaller, quicker, cheaper (SQC) methods of literacy assessment" that are "just big enough, faster at capturing and analysing data and

cheaper in terms of time and effort." This means that non-governmental organizations have been able to design and conduct hybrid assessments administered through household surveys that capture data from children in and out of school. Nonetheless, some of these surveys, such as *Uwezo*, draw on funding and expertise from international agencies and are influenced by their policy agendas (Languille 2014). Nonetheless, the simplicity of the assessments enhances their communicability at the national and sub-national level and makes them amenable to use by local stakeholders to hold national and local governments to account for learning outcomes.

64. Hyrbid assessments tend to collect less data on school and student characteristics and cannot compare policy and other national level systemic factors. They are less well suited for exploring associations between learning outcomes and other variables so care needs to be taken in drawing implications from these studies on quality. Judgements on school quality need to also take into account the baseline of literacy and language skills with which learners enter school and the home environment of students. In contexts where teachers have been the subject of a discourse of derision and blamed for poor quality education, hybrid assessments may perpetuate this perception because they do not offer information on how well teachers are supported and enabled by education systems. Hybrid assessments may be effective in highlighting where there are serious issues of poor quality and inequity to be addressed in primary education but do not, on their own, provide enough information to identify the underlying causes. The more high profile hybrid assessments that have been conducted focus on literacy and numeracy; they do not address relevance or the ESD and decent work targets.

65. Learning surveys, in general, typically focus on a smaller number of competencies than national assessments. International and regional LSEAs and hybrd assessments have been used extensively to measure competencies in literacy and mathematics. IEA, through its Trends in Mathematics and Science Study (TIMSS) and its International Civics and Citizenship Study (ICCS), does evaluate learning outcomes in Science and Citizenship. Regional surveys have the option to design evaluations to areas of specific interest. So, although there is scope to assess targeted sets of knowledge and attitudes within the ESD and life skills domains, LSEAs cannot be as comprehensive as national assessments.

66. In conclusion, there are various candidate measures with different strengths and weaknesses available to be used as indicators of learning outcomes, particularly for literacy and numeracy. Rose (in press) argues that global monitoring should, and already does within the EFA GMR, draw on all of these. Over-specifying indicators by selecting just one measure of learning will weaken relevance and fitness-for-purpose. It is always possible to devise more complex and statistically robust international LSEAs and this will always be of interest to international experts and international agencies and will enable more detailed reporting on learning outcomes at the global level. Extending technologies of international assessment in the current global policy context may also be expected

to consolidate the architecture of global governance, extending its influence over low and middle income countries. However, the prominence of concerns around learning outcomes over the last five years (Center for Universal Basic Education at Brookings 2011; UNICEF/UNESCO 2013) is testament to the sufficiency of what already exists is to highlight inequalities in learning at the global scale. The ongoing global monitoring of learning outcomes is undeniably valuable for identifying and addressing inequalities between countries and cross-national trends but there are diminishing returns in extending this into the multiple and complex domains of learning that contribute to sustainable development, peaceful societies and skills for decent life and work in diverse contexts. Investing in assessment at the national level and the curricula that define the competencies against which students are assessed will allow for monitoring of a broader range of outcomes.

4c. Indicators for equitable outcomes and relevant learning

67. Two highly respected commentators set out opposing arguments regarding the monitoring of learning outcomes in EFA. The former Director of the EFA GMR, Pauline Rose (2015) argues that measures of literacy and numeracy are readily measurable and comprehensible and hence have considerable traction. Further, literacy and numeracy are important foundational skills that enable learning in other areas as well as participation in society. She concludes they are the most suitable indicators for monitoring how well we are doing at providing a quality education for the most marginalized.

68. Robin Alexander (2015), by contrast, presents the view that education quality centrally concerns what happens in the classroom, the interactions between teachers and students that constitute pedagogy. From the point of view of someone who has spent a lifetime researching schools and classrooms, standardized tests – particularly using test instruments largely designed at the international level – seem inadequate instruments for indicating quality (Alexander 2000). He stops short of saying that classroom processes should be monitored globally but rather questions whether the level of the learner is where global monitoring should focus (Alexander 2008). Looking only at learning outcomes, he argues, focuses attention on individuals, so that schools, teachers, learners themselves and their families are blamed for poor learning outcomes. Indeed, this kind of rhetoric or "discourse of derision" is often deployed by policy makers to deflect attention from systemic failings. As argued in Section 3b, system level indicators can serve to focus attention on the resources and conditions that enable teachers and learners. However, indicators need to be formulated with care to ensure that they support the right to education and respect cultural differences between education systems.

69. Figure 5 draws on both of these positions to recommend a set of indicators that include outcomes indicators of equity, system level process indicators for equity and relevance, and learner level indicators for the inputs needed for learner readiness. It does not address classroom processes, with which Alexander is concerned, because these have been bracketed into the teachers' target and are discussed in section 5 below.

70. Experience with the EFA goals suggests that system level process indicators will have less traction than equitable outcomes and inputs indicators, although all three are needed to ensure equitable and relevant outcomes. Their interpretation and implementation is subject to professional judgment. This could work in one of two ways. Indicators could be devised and defined at the international level and re-contextualised for the national level by international experts working with high-level national educational officers and then communicated to lower level educational officers and teachers tasked with their implementation. Alternatively, indicators could be used in a more open way to stimulate professional debate across levels. Professional debate then becomes the means for developing a shared understanding of terms such as "appropriate to age," "responsive to learners' needs" and "fair assessment" and the means for developing expertise for implementing the indicators. Whilst analysis of data may demand expertise, system level data is generally readily available, particularly in countries with one national curriculum and one national assessment board. Nonetheless, care would have to be taken to establish mechanisms for monitoring and reporting that are not onerous for governments; low income countries already expend considerable resources engaging with a large number of development partners.

71. The learner level inputs indicator, in particular, lends itself to being modified according to schooling conditions in different contexts. For example, chairs and desks considered essential in schools in much of the world are not used in many Indian primary schools (Alexander 2000). Not owning a school uniform is a reason for exclusion of Maasai girls living in chronic poverty in a remote part of Tanzania (Raymond 2014) but not an issue in many wealthier countries. Indicator (3) could be adapted by explicitly identifying benchmarks for the most marginalized groups within a country. The third learner level indicator may be addressed through existing LSEAs that already collect data on student characteristics, using questionnaires to ask questions such as how many meals students eat a day.

FIGURE 5. INDICATORS FOR BASIC EDUCATION TARGET

Equity in educational outcomes

- Proficiency in literacy and mathematics at various points in the basic education cycle as measured by **existing** surveys, including international and regional large-scale educational assessments and hybrid assessments. Disaggregated according to gender, socio-economic status, rural/urban location and types of special needs.
  Monitored at international level
- Performance in national examinations, disaggregated by gender, school location and special needs/disabilities.
- Transition rates to different forms of post-basic education, disaggregated by gender, school location and, if available, information on special needs.|
  Variables for disaggregation determined at national level to be responsive to available data and patterns of marginalization

System level processes to ensure equity and relevance

- Curriculum and assessment is appropriate to learners' age and level of schooling and responsive to their learning needs, with particular attention to the most marginalized groups.
  In some countries, this may need to be elaborated at sub-national level to be responsive to learners' language capabilities and the prevalence of printed literature in different communities
- National assessments are fair, robust and transparent.
  Involves professional judgment and international sharing of assessment expertise
- National learning objectives are linked to national economic, social and environmental development priorities and lay a foundation for developing knowledge, skills and values for decent life and work across the economic, social and environmental contexts within which learners live.
  Needs to be elaborated at national level
- National learning objectives are the subject of informed public and professional debate and consultation, and data relating to each indicator is disseminated to stakeholders including the general public, parents and educational professionals through accessible formats.
  Appropriate means of dissemination to be determined at national and sub-national level but should include use of independent public media

Learner level inputs to ensure readiness to learn

- All learners are ready to learn, including not being hungry, having access to safe drinking water at school and basic equipment for learning, such as pens, exercise books and textbooks that are necessary for learning, with particular attention to the most marginalized.
  Inputs to be selected at national and/or sub-national level to be responsive to the most prevalent forms of need

5. INDICATORS FOR TEACHERS AND EDUCATIONAL PROCESSES

72. In the lead up to 2015, two different genres of literature have highlighted the importance of teachers to education quality – reports and research written or commissioned by rights-based advocacy organizations and large scale comparative studies oriented towards informing policy.

5a. Teachers within the rights-based tradition

73. Proposals for a target on teachers as part of an education SDG address a key concern of rights-based organizations with educational processes and outcomes. The inclusion of quality targets is also consistent with comprehensive understandings of quality in the Dakar Framework and Jomtien Declaration. However, this is the first time an attempt has been made to back quality targets with measurable indicators (Rose 2015) and there is still some way to go to formulate indicators. The EFA Steering Committee TAG (2014) only suggested input indicators for teachers, namely the percentage qualified to national standards and the percentage with pedagogical training. It also recommended disaggregation by gender. UNESCO Institute of Statistics already collects data on teacher qualifications and the latest EFA GMR (UNESCO 2014) highlights the challenge teacher supply presents for expanding enrolments rapidly from a low base. Teacher supply creates an intergenerational link between the quality of education systems present and future, and thus relates to sustainability. Expanding an education system faster than its capacity to supply teachers can depress quality, as seen in the first decade of universal primary education in Malawi and Uganda (Lewin 2009), postponing the point at which universal access is achieved. A teacher target therefore may work to slow down the rate of educational expansion, ensuring that expansion is achieved alongside quality improvement. Ultimately, this means universal basic education is achieved more quickly and efficiently due to lower rates of repetition and drop-out.

74. The Global Campaign for Education (GCE) has made teachers one of its main themes and in a proposal for post-2015 targets and indicators made teachers central to quality (GCE 2014). This document proposes a target that not only specifies an expectation for teachers to be qualified but articulates criteria for teacher training:

> By 2025, all children are taught by qualified teachers who have training in pedagogy, rights and gender sensitivity, in an accessible and safe environment. (GCE 2014: 3)

75. There are two suggested indicators for the teacher component of this target, the first of which combines a quantitative input with qualitative conduct components:

> Percentage of children taught by trained and qualified teachers, with clear and transparent national benchmarks for qualified teacher status which includes training in pedagogy, rights and gender sensitivity. (GCE 2014: 3)

The wording suggests that GCE has anticipated that teacher supply may be increased by reducing the rigour of training (for example, by lowering entry requirements or

shortening the duration of training). The conduct component calls for the exercise of professional judgment. How that judgment is exercised, and by who, is critical in determining how such an indicator would work in practice to improve quality. Much depends on establishing a shared understanding of benchmarks for pedagogy, rights and gender sensitivity. As discussed above with respect to system level learning indicators, professional debate has the potential to develop and disseminate understandings of complex subjective concepts. It matters, therefore, that debate is conducted not just amongst international experts meeting in metropolitan centres but also at different levels within education systems, including within teacher education institutions. The post-2015 successor to the GMR could catalyze debate through including discursive sections for reporting on conduct indicators, which report on their interpretation as well as their implementation within different education systems.

76. GCE, Muscat and OWG all limit their targets to a concern with teachers as input, assuming an unproblematic relationship between training and the practice of qualified teachers; this is not borne out by education research. Small scale qualitative research looking at the practice of what are often rather loosely termed "learner-centred practices" across diverse contexts has shown how the efforts of teachers, particularly newly trained teachers, to implement change is constrained by various factors, including their own preconceptions about teaching and learning, the "hidden curriculum" of teacher education colleges, the school environment (large class size, absence of materials or little preparation time), the conservatism of more experienced and powerful staff, school culture and unreformed inspection practices (Lewin & Stuart 2003; Vavrus 2009; Mtika & Gates 2010; Sriprakash 2010; Schweisfurth 2013). In short, teacher training does not work on its own to change practices and needs to be backed by appropriate resources and school-based supervision to support teachers to change their practice. Professionalisation, as Johnson *et al.* (2000: 190) observe, "is essentially a systemic issue rather than an individual one." However, the nearest that the post-2015 proposals come to recognizing this is Muscat's reference to "well-supported" teachers (Global EFA Meeting 2014: 3).

77. It is worth noting that the phrase "teacher training" is used in the post-2015 proposals rather than "teacher education." The latter is taken to signify that teaching is a profession, entrance to which is dependent on higher education qualifications. In many low-income countries, qualified teacher status is a tertiary level qualification at a lower academic level than a university degree, delivered through specialist teacher colleges. The use of the phrase "training" may be construed as a signal that the teachers' target is directed towards low rather than high income countries. "Training" can further be interpreted as suggesting teacher preparation involves acquiring technical skills rather than professional expertise and judgment. This view of teacher preparation, however, overlooks the complexity of teaching as a socio-culturally embedded activity that involves engagement with children and young people from diverse backgrounds and with diverse abilities.

78. Using the term "professional development" rather than "training" opens up pathways for implementing the teachers' target, which capitalizes on knowledge and expertise within the body of people who work as teachers. Professional development is ongoing and can be more or less formal, ranging from collegial interactions and mentoring relationships to university degrees. Whilst training is typically a self-contained systematic programme of activity rolled out by a government institution or external body, professional development may be developed and led by practicing teachers within or outside of schools. Professional associations, for example of subject specialist teachers, bring together individuals with the greatest enthusiasm for extending knowledge and expertise in teaching; these stakeholders can be agentic in developing ideas and disseminating them through the teaching population. So whilst "professional teachers" may seem a higher bar to achieve than "trained teachers" it distributes the responsibility away from centres of administration to the teaching body as a whole, drawing attention to the most widely available human resource for improving teaching and learning, namely teachers themselves (Samoff *et al.* 2011).

79. Some research by rights-based organizations has looked at the living and working conditions of teachers. VSO (2002), for example, studied teacher motivation across three countries. The research on teachers' living conditions and salaries (e.g. Marphatia *et al.* 2007) is an important strand to this work, highlighting teachers' working and living conditions as human rights issues in themselves, whilst having profound implications for educational quality. Studies remain relatively small scale compared to LSEAs (see below) and whilst they identify and expand understanding of a key issue for international research, more work is needed to identify potential data sources. System level data may be fairly straightforward to collect, but analysis depends on professional judgment. What levels of renumeration mean in practice for the lifestyles and livelihoods of teachers serving in diverse contexts and their professional conduct and practices, however, is harder to ascertain and may continue to be the subject of small scale qualitative research (e.g. Tanaka 2010; Buckler 2014; Tao 2014). Policy research on teachers, by comparison, particularly in OECD countries, is already capturing data on teachers and proposing benchmarks.

5b. Teachers within policy research

80. The other set of literature that places a primacy on teachers as determinants of education quality is large-scale policy-oriented research. Since the beginning of the 2000s, the teaching profession has become a focus for global policy debate and, arguably, the emphasis on the role of teachers for educational reform has never been more pronounced than now (Connell 2009; Robertson 2012). OECD with the programme *Teaching and Learning International Survey* (TALIS), the World Bank with SABER-*Teachers*, and UNESCO are the main policy actors on the international level in this arena, each with their distinctive profile. OECD's large-scale international survey TALIS is the most comprehensive in terms of collecting detailed cross-national data (Robertson 2012; Robertson 2013). We briefly review this study in order to identity the kind of data that could

potentially become available for a broader range of countries over the next 15 years. Two rounds of the survey have been conducted so far, in 2008 and 2013, with 24 and 34 countries or regions participating respectively. It is anticipated that future rounds will include more countries, OECD members as well as non-members.

81. TALIS basically consists of two questionnaires, one for teachers and one for school heads in lower secondary schools. In addition to this primary target group, for TALIS 2013 participant countries were given the option to include primary and upper secondary school teachers and heads, and to link TALIS data to PISA. Linking the two data sets makes it possible to explore associations between TALIS variables and student learning outcomes. The policy foci in TALIS are determined by a joint priority-rating exercise by participating OECD member countries (see figure 6).

82. What is immediately apparent is that TALIS draws on the school effectiveness narrative that it shares with PISA. This entails that TALIS puts the teaching profession and the quality of teaching at the crux of education reform and economic growth. The key argument is that teachers as "the front-line workers" play a crucial role in the modernization of education systems because, within schools, "teacher- and teaching-related factors are the most important factors that influence student learning" (OECD 2014: 32).

83. At the same time, TALIS serves the purpose of bringing "the voice of teachers" into the debate. Symptomatically, the global teacher union Education International has endorsed and been involved in the survey programme. However, it would seem that the voice of teachers is circumscribed and framed by the overarching narrative of school effectiveness and student performance as assessed in PISA. This means that some of the critical points concerning insensitivity towards the distinctiveness of socio-cultural environments can also be raised against TALIS. On this basis, Sobe (2013) associates TALIS with the construction of a simplistic and reductive "global reality of teacher professionalism" driven by standardization, codification and identification of educational "best practices."

FIGURE 6. TALIS POLICY THEMES

TALIS 2008 – three main themes	TALIS 2013 – five main themes
School leadership	School leadership, including new indicators on distributed or team leadership
Appraisal of and feedback to teachers	Appraisal of and feedback to teachers
Teaching practices, beliefs and attitudes	Teachers' pedagogical beliefs, attitudes and teaching practices, including new indicators on the profile of student assessment practices
+ Professional development of teachers as "an important theme" due to synergies with three main themes and European Union interest	Teacher training, including professional development and new indicators on initial teacher education
+ Aspects of other themes: school climate, division of working time and job satisfaction	Teachers' reported feelings of self-efficacy, their job satisfaction and the climate in the schools and classrooms in which they work

(OECD 2009; OECD 2014)

84. In relation to the questions that we should ask of indicators of quality, TALIS is clearly a major research exercise based on a sophisticated conceptual framework. TALIS thus offers a wealth of insights into the state and nature of the teaching profession and teachers' thinking of their practices and status in society. In terms of putting a focus on teachers' work conditions, self-efficacy and status, TALIS indicators have much to offer. However, how well TALIS is compatible with a rights-based view of education remains a question. Since TALIS has so far only been conducted twice, we know little about how it contributes towards improving education quality. The involvement of civil society, represented by teacher unions and business organizations, in the conception and implementation of TALIS is a positive. Yet, overall the impact and direction of the programme is likely to be determined by its relationship to PISA in future rounds.

85. The McKinsey reports (Barber & Mourshed 2007; Mourshed *et al.* 2010) and "Learning Curve" study (The Economist Intelligence Unit 2014), discussed earlier, have also surfed on the wave of global political attention directed towards the teaching profession. Indeed, they have arguably been instrumental in directing political gaze towards teachers. Private consultancies and corporate philanthropists have gained considerable political influence despite having little in-house educational expertise (Ball 2012). As mentioned above, their research has been critiqued by scholars for promoting one size fits all solutions that do not recognize the cultural and political specificity of education systems (Coffield 2012; Crossley 2014; Morris in press). Some have further observed the financial or political interest these organizations, which operate beyond national spaces of representation and democratic accountability, have in expanding their influence in global education governance (Ball 2012; Robertson *et al.* 2012). What is worth noting is the presence and influence of these organizations within the field of education internationally, their active promotion of the learning outcomes agenda, their involvement in policy research scrutinizing the role of teachers, and their participation in the post-2015 education debates (Robertson 2012; McLean 2013).

86. Regional LSEAs can also and in some cases do collect data from teachers. SACMEQ includes a teacher questionnaire for Grade 6 teachers of literacy and mathematics, and also administers the pupil tests to teachers. The teacher questionnaire elicits information on classroom processes, teacher qualifications and job satisfaction. However, associations between teaching processes and learning outcomes are elusive and tend to be weak. There are several reasons why this may be. Some analyses suggest that for schools serving low socio-economic groups, resources have a stronger association with learning outcomes than process variables (Smith 2011). In any context, children's learning outcomes are influenced by teaching processes encountered over their whole school career and surveys only capture data from their current teacher. Fine-grained qualitative research on teachers shows that teachers' espoused values and the claims they make about their practice can diverge from observed practices (e.g. Osborn *et al.* 2000; Schweisfurth 2002). This is a reminder of Alexander's (2008) assertion that it is not teachers *per se* that lie at the heart of quality but what they *do*. Teachers' responses to questionnaires, therefore, like learners' performance in standardized tests, should be regarded as a proxy indicator that gives only partial information on quality. Despite their growing sophistication and yield of large complex data sets, large-scale surveys remain blunt instruments for researching teaching and learning processes.

5c. Indicators for professional teachers and quality teaching processes

87. It is neither feasible nor desirable to monitor classroom processes at the global level. This is properly the work of the system level, carried out through its inspectorate and other forms of school supervision, supplemented by internal evaluation processes such as school self-evaluation (Carlson 2009). There are also examples of community involvement in assessing the quality of schools and holding teachers to account for school quality (Prew & Quaigrain 2010). Global monitoring could focus on the system level functions of regulating and monitoring quality. Together with curriculum and assessment, this would directly address the terms by which teachers' work is judged and evaluated. A school supervision indicator would address the "well-supported teachers" component of the Muscat target. Research evidence suggests that, together with poor renumeration and perceived lack of social status, neglect by over-stretched district authorities is a cause of demoralisation amongst teachers serving in rural schools in poorly resourced systems (VSO 2002; Barrett 2005; Mpokosa & Ndarahutse 2008). A community of teachers working in difficult conditions with the minimum of teaching and learning resources can slip into poor professional or, sometimes, unprofessional and unethical practices (Anangisye & Barrett 2005). They are also denied the rewards of a career path where access to in-service training and promotion or transfer to another school is controlled by local education officials. In some countries it is not uncommon for local education offices and/or school inspectors to be so stripped of funding that they have no vehicles or petrol for vehicles (de Grauwe 2001). However, there are different models of school supervision. These have evolved over time, are embedded within the structures of education systems, are contingent on assumptions and models of teacher education, and work in concert with internal and local school evaluation processes (de Grauwe 2008). An indicator for

global monitoring would need to be open enough to accommodate these diverse models without imposing standardization. This could be achieved by focusing on the criteria used to evaluate education quality rather than the structures and forms of supervision.

88. We suggest, therefore, a pair of indicators that mimic the structure of the GCE proposal for a teachers' indicator by matching a quantitative input indicator that sets expectations for the frequency of contacts between a school and its supervisors together with a qualitative process indicator that sets expectations for the criteria use to evaluate schools. The latter would need to be elaborated at the national level.

89. Education systems can collect robust data on frequency of school visits, evaluations or inspections, which can be readily disaggregated by school location. Such an indicator may work to ensure that distribution of resources for school supervision units takes into account the logistic and other challenges of the locality they serve. So, for example, inspectors serving rural districts with dispersed populations have motorbikes and funds for fuel and inspectors can spend more time at schools needing more external professional support. The second indicator is not readily measurable or robust but intended to open up scrutiny and debate over the criteria by which teachers' work is judged. It extends beyond classroom teaching and learning to school processes more generally, including the hidden or informal curriculum that regulates student behavior, school environment and resources, and the participation of students and local community members in school decision-making processes. Theorists have highlighted these as being essential domains of concern for both the right to education and education for sustainable development but well beyond the focus of the learning outcomes agenda. The indicator has deliberately been formulated to refrain from prescribing specific criteria, such as, for example, not using corporal punishment and humiliation. This level of detailed prescription is the remit of national or sub-national level decision-making. Preserving the space for national level decision-making allows for national ownership, preserves the space for debate and decision making that also develop capacity, and allows for criteria to be expressed in terms that have contextual relevance and socio-cultural significance.

90. The effectiveness of school inspections or evaluations to improve quality depends on the information being communicated in a clear way to teachers and the communities served by schools. A third related indicator could address how information on school quality is disseminated and used.

91. Of course, many countries already have functioning supervision systems and expectations for professional development that have been developed over decades drawing on extensive professional experience. Beyond accountability tools for evaluating education systems, indicators must also be used to start conversations between and within countries through which expertise and knowledge is shared internationally.

FIGURE 7. INDICATORS FOR PROFESSIONAL TEACHERS AND EDUCATIONAL PROCESSES

Learner level input for equitable quality education

- Every learner is in a class with a teacher to learner ratio more than one to fifty.
- Every learner has a teacher who engages in regular continuing professional development and is supervised on a regular basis.

Inputs for professional teachers and schools

- All schools are visited at least once a month during term time by an external supervisor, with particular attention to schools in rural and remote locations.
- Teachers engage in at least one day of formal professional development for every year that they teach.
  Figures should be interpreted as lowest possible level and countries can set higher levels, which may represent current practices or an ambition for improving inputs for teacher professionalism.
- Teachers' remuneration, living and working conditions meet the criteria of decent life and work

System processes that ensure quality of school and classroom processes

- Schools are evaluated according to transparent criteria consistent with the right to education and with the principles of education for sustainable development.
- Teacher education and professional development promotes the right to education and education for sustainable development.
- Teachers have the freedom and resources to form professional associations for the purpose of improving teaching and learning, maintaining ethical standards and protecting their rights as employees.
- Teachers are drawn from all sections of society, including the most marginalized groups.

School level processess

- Information on school quality, such as inspection or evaluation reports and assessment data, are made publicly available and accessible to parents and the local community.
- Representatives of parents and local community participate in decision making in all aspects of school life.
  Mechanisms and institutions for this are set at national level

6. CONCLUSION: RE-VISIONING GLOBAL MONITORING

92. The discussion in this paper is marked by a wariness of over-using technologies of quantification that assume the authority of objectivity or neutral measurement at the global level. Deployed within a context where global actors are able to gain increasing influence over education governance, indicators that are narrowly concerned with quantifiable measures of learning can close off possibilities for education quality and pathways for improving educational processes. A key question for us is whether it is possible to formulate indicators for the OWG and Muscat targets that transform global monitoring into a process that supports and enables stakeholders at all levels to find ways to transform educational processes.

93. Human rights treaties and 25 years of work exploring what they mean for education quality have already created a bold and broad agenda for education quality and equity. But it is an agenda that is at risk of being cropped "back to basics" by the current preoccupation with learning outcomes. The sustainable development agenda demands an expansion of the broad rights-based agenda for education, viewing processes as continuous with outcomes. So, for example, preparing learners to establish peaceful societies involves learners being actively engaged in analyzing peace and violence within their environments and trying to manage conflict non-violently. Learning to contribute towards environmental sustainability involves engaging with individuals and institutions that are striving to better understand the natural environment and find ways of pursuing decent life and work sustainably. Both Muscat and OWG include targets that point us towards this broader vision without elaborating on how it can be achieved. Believing that education processes and outcomes are situated, we have tried to formulate indicators that encourage problem solving within and across education systems (see figures 5 and 7). However, these have the status of suggestions, which can be taken further.

94. The indicators we suggest for relevant learning include quantitative outcomes indicators but also qualitative process indicators intended to stimulate debate around what is relevant learning for different contexts and hence lead to the formulation of national and sub-national indicators that support and enable equitable, relevant learning for sustainable development. The indicators for teachers draw on the notion of teacher professionalism to challenge treatment of teachers as inputs to be passively moulded by training. Instead, they are viewed as active collaborators in creating and implementing a broad and bold agenda that responds to the needs and contexts of their students.

95. Finally, we recognize that no set of indicators can ensure quality and equity. It therefore matters how indicators are implemented. A more complex agenda, particularly one that includes process indicators that are subject to professional judgment, opens up opportunities for international technical experts, well-versed in the reasoning that emanates from the world's cosmopolitan centres, to re-contextualize indicators for

national and local actors, who then have a limited sense of ownership of the indicators. To use indicators supportively within a more collaborative process of creating and implementing new understandings of education quality demands relentless reflexivity. Such reflexivity is too often squeezed out within the culture of performativity and results-based management that dominates much of the development field internationally. Just as education systems should create the conditions for teachers as professionals to be active in creating as well as implementing a vision for education for sustainable development, so global monitoring should be aimed at creating the conditions that support policy-makers and education professionals to contribute towards an expansive and expanding agenda for quality and equity.

REFERENCES

Ahmed, M. (2014). Squaring the circle: EFA in the post-2015 global agenda, *International Journal of Educational Development,* 39: 64-69.

Alexander, R.J., (2000). *Culture and pedagogy: International comparisons in primary education.* Oxford: Blackwell.

Alexander, R.J. (2008). *Education for All, the quality imperative and the problem of pedagogy,* CREATE Research Monograph no. 20. Brighton: University of Sussex.

Alexander, R.J. (2015). Teaching and learning for all: the quality imperative revisited, *International Journal of Educational Development,* 40: 250-258.

Anagnostopoulos, D. (2007). The new accountability and teachers' work in urban high schools in USA. In Tatto, M. (ed). *Reforming Teaching Globally.* Oxford: Symposium, 119-135.

Anangisye, W.A.L. & Barrett, A.M. (2005). Professional identity and misconduct: perspectives of Tanzanian teachers. *Southern African Review of Education with Education with Production,* 11: 5-22.

Ball, S. J. (2012). *Global Education Inc: New Policy Networks and the Neo-liberal Imaginary.* London: Routledge.

Barber, M., & M. Mourshed (2007). *How the World's Best-Performing School Systems Come Out on Top.* McKinsey & Company.

Barrett, A. M. (2005). Teacher accountability in context: Tanzanian primary school teachers' perceptions of local community and educational administration. *Compare: a Journal of Comparative and International Education,* 35(1): 43-61.

Barrett, A. M. (2011a). A Millennium Learning Goal for education post-2015: a question of outcomes or processes. *Comparative Education* 47(1): 119-133.

Barrett, A. M. (2011b). An education Millennium Development Goal for quality: complexity and democracy. *Compare: a Journal of Comparative and International Education,* 41(1): 145-148.

Barrett, A. M. (2013a). Education and other sustainable development goals: a shifting agenda for comparative education, Compare Forum. *Compare: a Journal of Comparative and International Education,* 43(6): 825-829.

Barrett, A. M. (2013b). Measuring learning post-2015, paper presented at the 12th UKFIET International Conference on Education and Development, 13-15 September 2013. http://research-information.bristol.ac.uk/files/15034109/AMBarrett_UKFIET_2013_Measruing_learning_post_2015.pdf, last accessed 16 April 2015.

Barrett, A. M. (2014). Measuring literacy post-2015: some social justice issues. In McIlwraith, H. (ed). *Language Rich Africa Policy Dialogue: The Cape Town Language and Development Conference: Looking beyond 2015.* London: British Council, 71-77.

Beyond2015 (2013). *Making Education For All a Reality, Global Thematic Consultation on Education and the Post-2015 Development Framework.* March 2013. URL: http://www.beyond2015.org/sites/default/files/Beyond%202015-Education.pdf, last accessed 10 April 2015.

Bloem, S. (2013). *PISA in low and middle income countries.* OECD Education Working Paper No. 93, OECD.

Bloem, S. (in press). PISA for low and middle income countries, Compare Forum. *Compare: a Journal of Comparative and International Education.*

Carlson, B. (2009). School self-evaluation and the 'critical friend' perspective. *Educational Research and Review* 4(3): 78-85.

Carr-Hill, R. (2012). Finding and the counting out-of-school children. *Compare: a Journal of Comparative and International,* 42(2): 187-212.

Center for Universal Education at Brookings. (2011). *A global compact on learning: taking action on education in developing countries.* Washington DC: Center for Universal Education at Brookings.

Chung, J. (in press). International comparison and education policy learning: looking North to Finland, Compare Forum. *Compare: a Journal of Comparative and International Education.*

Coffield, F. (2012). Why the McKinsey reports will not improve school systems, *Journal of Education Policy,* 27(1): 131-149.

Commonwealth Ministerial Working Group on the Post-2015 Development Framework for Education. (2012). *Commonwealth Recommendations for the Post-2015 Development Framework for Education, Background Paper.* London: Commonwealth Secretariat.

Connell, R. (2009). Good teachers on dangerous ground: Towards a new view of teacher quality and professionalism. *Critical Studies in Education,* 50(3): 213-229.

Croft, A. (2006). Prepared for diversity? Teacher education for lower primary classes in Malawi. In Little, A.W. (ed). *Education for All and Multigrade Teaching: Challenges and Opportunities.* Dordrecht: Springer, 103-126.

Crossley, M. (2014). Global league tables, big data and the international transfer of educational research modalities. *Comparative Education* 50(1): 15-26.

Dahler-Larsen, P. (2012). Constitutive Effects as a Social Accomplishment: A Qualitative Study of the Political in Testing. *Education Inquiry,* 3(2): 171-186.

Davidson, E. (2004). *The progress of the Primary Education Development Plan in Tanzania: 2002-2004.* HakiElimu Working Paper 04.2. Dar es Salaam: HakiElimu.

De Grauwe, A. (2001). *School supervision in four African countries, Vol 1. Challenges and reforms.* Paris: International Institute for Educational Planning/UNESCO. http://www.africanafrican.com/folder15/alot%20more%20of%20african%20%26%20african%20american%20history13/botswana/124823e.pdf, last accessed 13 March 2015.

De Grauwe, A. (2008). *School Monitoring Systems and their Impact on Disparities, Background paper prepared for the Education for All Global Monitoring Report 2009 - Overcoming Inequality: why governance matters.* Paris: UNESCO. http://unesdoc.unesco.org/images/0018/001800/180083e.pdf, last accessed 13 March 2015.

Delors, J. *et al.* (1996) Learning the treasure within, Report to UNESCO of the International Commission on Education for the Twenty-first Century. Paris: UNESCO.

EFA GMR. (2013). *Proposed post-2015 education goals: Emphasizing equity, measurability and finance.* Initial draft for discussion, March 2013. Paris: UNESCO.

EFA GMR. (2014). *Concept note for a 2016 report on 'Education, sustainability and the post-2015 development agenda'.* Paris: UNESCO. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/GMR/images/2014/2016_Concept_Note_rev2.pdf, last accessed 11 March 2015.

EFA Steering Committee TAG on the post-2015 education indicators (2014). *Monitoring the post-2014 education targets: a note on indicators.* Paris: UNESCO.

Elgert, L. & Krueger, R. (2012). Modernising sustainable development? Standardisation, evidence and experts in local indicators. *Local Environment* 17(5): 561-571

Fien, J. & Tilbury, D. (2002). The global challenge of sustainability. In Tilbury, D., Stevenson, R.B., Fien, J. & Schruder, D. (eds). E*ducation and Sustainability: Responding to the Global Challenge.* Gland, Switzerland & Cambridge, UK: International Union for Conservation of Nature and Natural Resources (IUCN), 1-12.

Filmer, D., Hasan, A. & Pritchett, L. (2006). *A Millennium Learning Goal: Measuring real progress in education.* Working paper 97. Washington DC: Center for Global Development.

Fraser, N. (2008). *Scales of Justice: Reimagining Political Space in a Globalizing World.* Cambridge & Malden MA: Polity.

Fukuda-Parr, S., Yamin, A. E. & Greenstein, J. (2013). *Synthesis Paper - The power of numbers: A critical review of MDG targets for human development and human rights.* FXB Working Paper. Cambridge MA: Harvard School of Public Health/Harvard University/The New School.

GCE (2002). *A quality education for all: priority actions for governments, donors and civil society.* Brussels: Global Campaign for Education.

GCE (2014). *Equitable, inclusive and free: A collective vision for quality education beyond 2015.* Johannesburg: Global Campaign for Education. http://www.campaignforeducation.org/docs/post2015/GCE_POST2015_FINAL_GOAL_EN.pdf, last accessed 11 March 2015.

Global Education for All Meeting (2014). *2014 GEM Final Statement: The Muscat Agreement.* UNESCO.

Goldstein, H., (2004). International comparisons of student attainment: some issues arising from the PISA study. *Assessment in Education: Principles, Policy & Practice,* 11(3): 319-330.

Gove, A. & Wetterberg, A. (2011). The Early Grade Reading Assessment: An introduction. In Gove, A. & Wetterberg, A. (eds). *The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy.* North Carolina: RTI, 1-37.

Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy,* 24(1): 23-37.

Heneveld, W. & Craig, W. (1996). *Schools count: World Bank project designs and the quality of primary education in Sub-Saharan Africa.* Washington DC: The World Bank.

IEA. (2013). *International Civics and Citizenship Study 2016: A brochure.* IEA.

Johnson, S., Monk, M. & Hodges, M. (2000). Teacher development and change in South Africa: A critique of the appropriateness of transfer of northern/western practice. *Compare: a Journal of Comparative and International Education,* 30(2): 179-192.

Jolly, R., Emmerij, L. & Weiss, T. G. (2005). *The power of UN ideas: Lessons from the first 60 years. A summary of the books and findings from the United Nations Intellectual History Project.* New York: United Nations Intellectual History Project.

Kanjee, A. (2012). Options for developing integrated assessment systems to improve learning and teaching, Compare Forum. *Compare: a Journal of Comparative and International Education,* 42(3): 525-531.

Langford, M. (2012). *The art of the impossible: Measurement choices and the post-2015 development agenda.* Governance and human rights: Criteria and measurement proposals for a post-2015 development agenda, OHCHR/UNDP Expert Consultation. New York: UN Sustainable Development Knowledge Platform.

Languille, S. (2014). Quality education through performativity. 'Learning crisis' and technology of quantification in Tanzania. *International Journal of Educational Development,* 29: 49-58.

Learning Metrics Task Force. (2013). Towards universal learning: recommendations from the Learning Metrics Task Force. Montreal and Washington DC: UNESCO Institute of Statistics and Center for Universal Education at Brookings. September 2013.

Lewin, K. M. (2009). Access to education in sub-Saharan Africa: Patterns, problems and possibilities. *Comparative Education,* 45(2): 151-174.

Lewin, K. M. & Stuart, J. (2003). *Researching teacher education: New perspectives on practice, performance and policy, Multi-Site Teacher Education Research Project (MUSTER) Synthesis Report.* Brighton: University of Sussex.

Marphatia, A. A., Moussié, R., Ainger, A. & Archer, D. (2007). *Confronting the Contradictions: The IMF, wage bill caps and the case for teachers.* London: ActionAid.

Mpokosa, C. & Ndarahutse, S. (2008). *Managing Teachers: The Centrality of Teacher Management to Quality Education. Lessons from developing countries.* London: CfBT and VSO.

McCowan, T. (2010). Reframing the universal right to education. *Comparative Education,* 46(4): 509-525.

McCowan, T. (2013). *Education as a Human Right: Principles for a Universal Entitlement.* London: Bloomsbury.

McLean, H. (2013). Goldilocks revisited: Measuring education quality within the post-2015 global framework. *Commonwealth Education Partnerships* 2013/14: 64-68. http://www.cedol.org/wp-content/uploads/2013/11/1-Goldilocks-revisited.pdf, last accessed 13 March 2015.

Morris, P. (in press). Comparative Education, PISA, Politics And Educational Reform: A Cautionary Note, Compare Forum. *Compare: a Journal of Comparative and International Education.*

Mourshed, M., Chijioke, C. & M. Barber (2010). *How the world's most improved school systems keep getting better.* McKinsey & Company.

Mtika, P. & Gates, P. (2010). Developing learner-centred education among secondary trainee teachers in Malawi: The dilemma of appropriation and application. *International Journal of Educational Development* 30(4): 396-404.

Nardi, E., (2008). Cultural biases: A non-Anglophone perspective. *Assessment in Education: Principles, Policy & Practice,* 15(3): 259-266.

OECD. (2014). *TALIS 2013 Results: An international perspective on teaching and learning.* OECD.

O'Neill, O. (2002). *A Question of trust.* BBC Reith Lecture, 17 April 2002, URL: http://www.bbc.co.uk/programmes/poogpzdf, last accessed 28 February 2015.

Open Working Group for Sustainable Development Goals. (2014). *Proposal of the Open Working Group for Sustainable Development Goals.* New York: UN Sustainable Development Knowledge Platform.

Osborn, M., McNess, E., Broadfoot, P. with Pollard, A. & Trigg, P. (2000). *What Teachers Do: Changing Policy and Practice in Primary Education.* London: Continuum.

Palmer, R., Wedgwood, R., Hayman, R. with King, K. & Thin, N. (2007) Educating out of poverty? A synthesis report on Ghana, India, Kenya, Rwanda, Tanzania and South Africa. London: DFID.

Pigozzi, M. J. (2008). *Towards an index of quality education.* Paper prepared for the International Working Group on Education (IWGE).

Piper, B. & Miksc, E. (2011). Using Early Grade Reading Assessments to investigate language-of-instruction policy in East Africa. In Gove, A. & Wetterberg, A. (eds). *The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy.* North Carolina: RTI Press, 139-182.

Post-2015 Education Indicators TAG of the EFA Steering Committee (2014). *Towards indicators for a post-2015 education framework.* Montreal: UNESCO Institute of Statistics.

Prew, M. & Quaigrain, K. (2010). Using school performance data to drive school and education district office accountability and improvement: The case of Ghana. *Educational Management Administration & Leadership,* 38: 728-744.

Pritchett, L. & Beatty, A. (2015). Slow down, you're going too fast: Matching curricula to student skill levels. *International Journal of Educational Development* 40: 276-288.

Raymond, A. (2014). *Girls' education in pastoral communities: An ethnographic study of Monduli District, Tanzania.* London: CfBT. http://cdn.cfbt.com/~/media/cfbtcorporate/files/research/2014/r-girls-education-in-pastoral-communities-2014.pdf, last accessed 14 March 2015.

Ringarp, J. & Rothland, M. (2010). Is the grass always greener ...? The effect of PISA results on education debates in Sweden and Germany. *European Educational Research Journal* 9(3): 422-430.

Robertson, S. L. (2012). Placing Teachers in Global Governance Agendas. *Comparative Education Review,* 56(4): 584-607.

Robertson, S. L. (2013). Teachers' Work, Denationalisation, and Transformations in the Field of Symbolic Control: A Comparative Account. In Seddon, T. & Levin, J. (eds). *World Yearbook of Education 2013. Educators, Professionalism and Politics: Global Transitions, National Spaces and Professional Projects.* Oxford: Routledge. pp. 77-96.

Robertson, S. L., Mundy, K., Verger, A. & Menashy, F. (eds. 2012). *Public Private Partnerships in Education: New Actors and Modes of Governance in a Globalizing World.* Cheltenham: Edward Elgar.

Rose, P. (2015). Three levels for educational quality: Clarity, measurability and equity. *International Journal of Educational Development,* 40: 289-296.

Rose, P. (in press). Is a global system of international large-scale assessments necessary for tracking progress for a post-2015 learning target? Compare Forum. *Compare: a Journal of Comparative and International Education.*

Samoff, J., Dembélé, M. & Molapi Sebatane, E. (2013). Scaling up by focusing down: creating space and capacity to extend education reform in Africa. In Tikly, L. & Barrett, A.M. (eds). *Education Quality and Social Justice in the Global South: Challenges for Policy, Practice and Research.* London: Routledge, 121-138.

Schweisfurth, M. (2002). *Teachers, Democratisation and Educational Reform in Russia and South Africa.* Oxford: Symposium.

Schweisfurth, M. (2013). *Learner-centred Education in Perspective: Whose Pedagogy for whose Development?* London: Routledge.

Sellar, S. & Lingard, B. (2013). PISA and the expanding role of the OECD in global educational governance. In Meyer, H. D. & Benavot, A. (eds). *PISA, power, and policy: The emergence of global educational governance.* Oxford: Symposium, 185-206.

Sen, A. (2009). *The Idea of Justice.* London: Allen Lane.

Smith, M. C. (2011). Which in- and out-of-school factors explain variations in learning across different socio-economic groups? Findings from South Africa. *Comparative Education,* 47(1): 79-102.

Sobe, N. W. (2013). Teacher professionalization and the Globalization of Schooling. In Seddon, T. & Levin, J. (eds). *World Yearbook of Education 2013. Educators, Professionalism and Politics: Global Transitions, National Spaces and Professional Projects.* Oxford: Routledge, 42-54.

Sobhy, H. (2012). The de-facto privatisation of secondary education in Egypt: A study of private tutoring in technical and general schools. *Compare: a Journal of Comparative and International Education,* 42(1):47-67.

Sriprakash, A. (2012). *Pedagogies for Development: The politics and practice of child-centred education in India.* Dordrecht, Springer.

Srivastava, P. (ed.) (2013). *Low-fee Private Schooling: Aggravating Equity Or Mitigating Disadvantage?* Oxford: Symposium.

Subrahmanian, R. (2002). Engendering education: Prospects for a rights based approach to female education deprivation in India. In Molyneux, M. & Razavi, S. (eds). *Gender justice, development, and rights.* Oxford: Oxford University Press, pp. 204-237.

The Economist Intelligence Unit. (2014). *The Learning Curve: Education and Skills for Life, 2014 Report.* London: Pearson. URL: http://thelearningcurve.pearson.com/reports/the-learning-curve-report-2014, last accessed 14 March 2015.

Tikly, L. & Barrett, A. M. (2011). Social justice, capabilities and the quality of education in low income countries. *International Journal of Educational Development,* 31(1): 3-14.

UNESCO. (2004). *Education for All - The Quality Imperative: EFA Global Monitoring Report 2005.* Paris: UNESCO.

UNESCO. (2010). *Reaching the Marginalized: EFA Global Monitoring Report 2010.* Oxford/Paris: Oxford University Press/UNESCO.

UNESCO. (2013). The global learning crisis: why every child deserves a quality education. Paris: UNESCO. http://unesdoc.unesco.org/images/0022/002238/223826e.pdf, last accessed 17 April 2015.

UNESCO. (2014). *Teaching and Learning: Achieving Quality for All, EFA Global Monitoring Report 2013/4.*

Paris: UNESCO.

UNICEF. (2008). *Basic education and gender equality: quality of education.* http://www.unicef.org/girlseducation/index_quality.html, last accessed 22 January 2010.

UNICEF. (2009). *Manual: Child friendly schools.* New York: UNICEF.

UNICEF/UNESCO. (2013). *Making Education a Priority in the Post-2015 Development Agenda. Report of the Global Thematic Consultation on Education in the Post-2015 Development Agenda.* New York/ Paris: UNICEF/UNESCO.

Unterhalter, E. (2005). Global inequality, capabilities and social justice: the millennium development goal for gender equality in education. *International Journal of Educational Development*, 25(2): 111-122.

Unterhalter, E. (2014). Measuring education for the Millennium Development Goals: reflections for targets, indicators, and a post-2015 framework. *Journal of Human Development and Capabilities,* 15(2-3): 176-187.

Unterhalter, E. & Brighouse (2007). Distribution of what for social justice in education? The case of Education for All by 2015. In Walker, M. & Unterhalter, E. (eds). *Amartya Sen's Capability Approach and Social Justice in Education.* New York and London: Palgrave Macmillan.

Vavrus, F. (2009). The cultural politics of constructivist pedagogies: Teacher education reform in the United Republic of Tanzania. *International Journal of Educational Development,* 29(3): 3030-311.

Verger, A., Novelli, M. & Altinyelken, H. K. (2012). Global education policy and international development: An introductory framework. In Verger, A., Novelli, M. & Altinyelken, H. K. (eds). *Global Education Policy and International Development: New Agendas, Issues and Policies.* London: Bloomsbury.

Verger, A. & Novelli, M. (eds. 2012). *Campaigning for "Education for All": Histories, Strategies and Outcomes of Transnational Advocacy.* Rotterdam: Sense.

VSO. (2002). *What makes teachers tick? A policy research report on teachers' motivation in developing countries.* VSO: London.

Wagner, D. A. (2010). Quality of education, comparability, and assessment choice in developing countries. *Compare: A Journal of Comparative and International Education,* 40(6): 741-760.

World Bank. (2011). Learning for All: Investing in people's knowledge and skills to promote development. *World Bank Group Education Strategy 2020.* Washington DC: World Bank.

World Conference on Education for All. (1990). *World declaration on Education For All: meeting basic learning needs.* World Conference on Education for All, Jomtien, Thailand: UNESCO.

World Education Forum. (2000). *Dakar Framework for Action - Education For All: meeting our collective commitments.* World Education Forum, Dakar, Senegal: UNESCO.

APPENDIX 1. INDICATORS RELATED TO RELEVANT LEARNING OUTCOMES AND TEACHERS AND EDUCATIONAL PROCESSES

INDICATORS FOR RELEVANT LEARNING OUTCOMES

Equity in educational outcomes

- Proficiency in literacy and mathematics at various points in the basic education cycle as measured by **existing** surveys, including international and regional large-scale educational assessments and hybrid assessments. Disaggregated according to gender, socio-economic status, rural/urban location and types of special needs.
  Monitored at international level
- Performance in national examinations, disaggregated by gender, school location and special needs/disabilities
- Transition rates to different forms of post-basic education, disaggregated by gender, school location and, if available, information on special needs.
  Variables for disaggregation determined at national level to be responsive to available data and patterns of marginalization

System level processes to ensure equity and relevance

- Curriculum and assessment is appropriate to learners' age and level of schooling and responsive to their learning needs, with particular attention to the most marginalized groups.
  In some countries, this may need to be elaborated at sub-national level to be responsive to learners' language capabilities and the prevalence of printed literature in different communities
- National assessments are fair, robust and transparent.
  Involves professional judgment and international sharing of assessment expertise
- National learning objectives are linked to national economic, social and environmental development priorities and lay a foundation for developing knowledge, skills and values for decent life and work across the economic, social and environmental contexts within which learners live.
  Needs to be elaborated at national level
- National learning objectives are the subject of informed public and professional debate and consultation and data relating to each indicator is disseminated to stakeholders including the general public, parents and educational professionals through accessible formats.
  Appropriate means of dissemination to be determined at national and sub-national level but should include use of independent public media

Learner level inputs to ensure readiness to learn

- All learners are ready to learn, including not being hungry, having access to safe drinking water at school and basic equipment for learning, such as pens, exercise books and textbooks that are necessary for learning, with particular attention to the most marginalized.
  Inputs to be selected at national and/or sub-national level to be responsive to the most prevalent forms of need

INDICATORS FOR PROFESSIONAL TEACHERS AND EDUCATIONAL PROCESSES

Learner level input for equitable quality education

- Every learner is in a class with a teacher to learner ratio more than one to fifty.
- Every learner has a teacher who engages in regular continuing professional development and is supervised on a regular basis.

Inputs for professional teachers and schools

- All schools are visited at least once a month during term time by an external supervisor, with particular attention to schools in rural and remote locations.
- Teachers engage in at least one day of formal professional development for every year that they teach.
  Figures should be interpreted as lowest possible level and countries can set higher levels, which may represent current practices or an ambition for improving inputs for teacher professionalism.
- Teachers' remuneration, living and working conditions meet the criteria of decent life and work

System processes that ensure quality of school and classroom processes

- Schools are evaluated according to transparent criteria consistent with the right to education and with the principles of education for sustainable development.
- Teacher education and professional development promotes the right to education and education for sustainable development.
- Teachers have the freedom and resources to form professional associations for the purpose of improving teaching and learning, maintaining ethical standards and protecting their rights as employees.
- Teachers are drawn from all sections of society, including the most marginalized groups.

School level processess

- Information on school quality, such as inspection or evaluation reports and assessment data, are made publicly available and accessible to parents and the local community.
- Representatives of parents and local community participate in decision making in all aspects of school life.
  Mechanisms and institutions for this are set at national level